

DATE: Thursday, April 24, 2003 Printable Copy Create Case

---

### Search History

---

113	Search:
<input type="button" value="Refine Search"/>	<input type="button" value="Recall Text"/>
<input type="button" value="Clear"/>	
Database:	
US Patents Full-Text Database	
JPO Abstracts Database	
EPO Abstracts Database	
Derwent World Patents Index	
IBM Technical Disclosure Bulletins	

---

### Search Results -

0	19 and 15	0
<input type="button" value="Terms"/>	<input type="button" value="Documents"/>	

---

Main Menu | Search Form | Posting Counts | Show S Numbers | Edit S Numbers | Preferences | Cases

WEST

Set Name	Query	Slide by slide result set
L13	19 and 15	DB=USPT,PGBT; PLUR=YES; OP=AND
L12	19 and 14	0 L12
L11	19 and 13	0 L11
L10	19 and 11	0 L10
L9	monoephosphoryl adj lipid adj a or aluminum adj salt or qsl1 or isacadj 720 or saf or isocomes or mf-59 or sbas-2 or sbas-4 or detox or rc-529 or glucosamine-4-phosphate	15240 L9
L8	adjuvant and 13	0 L18
L7	adjuvant and 14	0 L17
L6	adjuvant and L5	0 L16
L5	5773246.pn.	1 L15
L4	5444149.pn.	1 L14
L3	5525495.pn.	1 L13
L2	adjuvant and L1	0 L12
L1	6255055.pn.	1 L1

#### END OF SEARCH HISTORY

L1	ANSWER 1 OF 36 MEDLINE Wang T; Fan L; Watanabe Y; McNeff J P D; Moulton G; Bangur C; Fangier G; R; Okada M; Inoue Y; Persling D H; Reed S G L523S, an RNA-biinding protein as a potential therapeutic target for lung cancer. JOURNAL code: 0370635. ISSN: 0007-0920.	AU
L4	ANSWER 2 OF 36 MEDLINE Mueller F; Bommer M; Lachter U; Ruhland C; Stagge V; Adler G; Gress T M; Seufertlein T KOC is a novel molecular indicator of malignancy. BRITISH JOURNAL OF CANCER, (2003 Mar 10) 88 (5) 699-701. Journal code: 0370635. ISSN: 0007-0920.	AU
L4	ANSWER 3 OF 36 MEDLINE Zhang Jian-Ying; Casiano Carlos A; Feng Xuan-Xian; Kozlak James A; Chan Edward K L; Tan Eng M Enhancement of Antibody Detection in Cancer Using Panel of Mack, David H.; Gish, Kurt C. Gene expression profiles useful for ovarian cancer and its use in diagnosis and screening for modulators of lung cancer and Murray, Natasha; Murray, Richard; Watson, Susan R.; Aziz, Natasha ANSWER 5 OF 36 CAPLUS COPYRIGHT 2003 ACS Aziz, Natasha; Murray, Richard Gene expression profiles associated with angiogenesis and their use in diagnosis and screening for modulators of lung cancer and PCT Int. Appl., 453 pp.	TI
L4	ANSWER 6 OF 36 CAPLUS COPYRIGHT 2003 ACS Murray, Natasha; Glynnne, Richard; Watson, Susan R.; Aziz, Natasha Gene expression profiles assessing for angiogenesis and their use in diagnosis and screening for modulators of lung cancer and PCT Int. Appl., 291 pp.	TI
L4	ANSWER 7 OF 36 CAPLUS COPYRIGHT 2003 ACS Mueller, Friederike; Gress, Thomas; Adler, Guido Medicament comprising a DNA sequence, which codes for the RNA-binding KOC protein, and comprising a KOC protein or a DNA sequence of the KOC promoter PCT Int. Appl., 75 pp.	IN

=> d au t1 so 1-36 14

FILE :MEDLINE, CAPLUS, BIOSIS, CISSEARCH, ENTERED AT 15:48:12 ON 24 APR 2003  
FILE :HOME, ENTERED AT 15:48:12 ON 24 APR 2003  
FILE :MEDLINE, CAPLUS, BIOSIS, CISSEARCH, ENTERED AT 15:48:26 ON 24 APR 2003  
1253 S KOC  
0 S MUELLER-PALLASCH AND LI  
71 S LI AND CANCER  
36 DUP REM 13 (35 DUPLICATES REMOVED)

- CODEN: PIXXXD2  
 ANSWER 8 OF 36 MEDLINE DUPLICATE 5  
 Hoo Linda SOO; Zhang Jianying Y.; Chan Edward K.L.  
 Cloning and characterization of a novel 90 kDa 'compantion' auto-antigen of p62 overexpressed in cancer x.  
 ONCOGENE, (2002 Jul 25) 21 (32) 5006-15.  
 journal code: 8711562. ISSN: 0950-9232.  
 ANSWER 9 OF 36 SCISEARCH COPYRIGHT 2003 ISI (R)  
 Iizamochi H (Reprint); Kigawa J; Sultana H; Iba T; Akeshima R; Kamazawa S;  
 Kanamori Y; Terakawa N  
 Sensitivity to anti-cancer agents and resistance mechanisms in clear cell carcinoma of the ovary  
 JAPANESE JOURNAL OF CANCER RESEARCH, (JUN 2002) VOL. 93, NO. 6, PP.  
 723-728.  
 Publisher: BUSINESS CENTER ACADEMIC SOCIETIES JAPAN, 5-16-9 HONKOMAGOME,  
 BUNKYO-KU, TOKYO, 113-8633, JAPAN.  
 ISSN: 0910-5050.  
 ANSWER 10 OF 36 CAPLUS COPYRIGHT 2003 ACS  
 Wang, Shanying; Liu, Long; Wang, Lin; Tian, Qizhou; Li, Qifu; Peng,  
 Xuanxian  
 Auto-antibody to Koc in patients with cancers  
 Zhaogguo Matanyixue Zazhi (2002), 18(1), 30-32  
 CODEN: ZMZAEE; ISSN: 1000-484X  
 ANSWER 11 OF 36 MEDLINE DUPLICATE 6  
 Kiyozuka Y; Nakagawa H; Senzaki H; Uemura Y; Adachi S; Teramoto Y;  
 Matsuyama T; Bessho K; Tsubura A  
 Bone morphogenic protein-2 and type IV collagen expression in psammoma body forming ovarian cancer.  
 ANTIcANCER RESEARCH, (2001 May-Jun) 21 (3B) 1723-30.  
 journal code: 8102988. ISSN: 0250-7005.  
 ANSWER 12 OF 36 BIOSIS COPYRIGHT 2003 BIOLGICAL ABSTRACTS INC.  
 Tan, Eng M. (1)  
 Autoantibodies as reporters identifying aberrant cellular mechanisms in tumors.  
 Martinez, Eduardo J.; Cory, B. J.; Owa, Takashi  
 Antitumor activity and gene expression-based profiling of ecteinascidin  
 CODEN: CBOLBZ; ISSN: 1074-5521  
 SO  
 ANSWER 13 OF 36 CAPLUS COPYRIGHT 2003 ACS  
 ISSN: 0021-9738.  
 Martinez, Eduardo J.; Cory, B. J.; Owa, Takashi  
 Antitumor activity and gene expression-based profiling of ecteinascidin  
 CODEN: CBOLBZ; ISSN: 1074-5521  
 SO  
 ANSWER 14 OF 36 BIOSIS COPYRIGHT 2003 BIOLGICAL ABSTRACTS INC.  
 Kohno, T. (1); Suzuki, M. (1); Mizukami, H.; Sagai, Y.; (1); Takei, Y. (1); Shiimpou, M.; Ohwada, M. (1); Kume, A.; Ozawa, K. (1); Sato, I. (1)  
 Expression of interleukin-10 inhibits angiogenesis and tumor growth in ovarian cancer.  
 European Journal of Cancer, (October, 2001) VOL. 37, NO. SUPPLEMENT 6, pp.  
 325. http://www.elesvier.com/locate/ejca. print.  
 Meeting info.: 11th European Cancer Conference Lisbon, Portugal October 21-25, 2001  
 ISSN: 0959-8049.  
 ANSWER 15 OF 36 BIOSIS COPYRIGHT 2003 BIOLGICAL ABSTRACTS INC.  
 Bommer, Martin (1); Seufferlein, Thomas; Gruss, Thomas; Adler, Guido;  
 Doebeke, Hartmut (1); Mueller, Friederike  
 AU  
 L4

TI	Detection of CNS involvement in patients with acute Leukemia and Non-Hodgkin's lymphoma using RT-PCR for Koc-expression.	Blood, (November 16, 2001) Vol. 98, No. 11 Part 2, pp. 157b.	http://www.bloodjournal.org/.print.	Meeting Info.: 43rd Annual Meeting of the American Society of Hematology, Part 2 Orlando, Florida, USA December 07-11, 2001	ISSN: 0006-4971.
AU	ANSWER 16 OF 36 BIOSIS COPYRIGHT 2003 BIOLGICAL ABSTRACTS INC.	Zhang, Jianying (1); Chan, Edward K.; Lu, MaoLiang; Wang, Xiao; Peng,	Xuanxian; Mueller-Pillasch, Friederike; Tan, Bing M.	Proceedings of the American Association for Cancer Research Annual Meeting, (March, 2001) Vol. 42, pp. 157. print.	Autotimune responses to mRNA binding proteins p62 and Koc in diverse malignancies.
AU	ANSWER 17 OF 36 MEDLINE COPYRIGHT 2003 ACS DUPLICATE 7	Zhang J Y; Chan E K; Peng X X; Lu M; Wang X; Mueller F; Tan E M	Zhang, Jianying (1); Chan, Edward K.; Lu, MaoLiang; Wang, Xiao; Peng,	Autotimune responses to mRNA binding proteins p62 and Koc in diverse malignancies.	Journal code: 100883537. ISSN: 1521-6616.
AU	ANSWER 17 OF 36 MEDLINE COPYRIGHT 2003 ACS DUPLICATE 7	Chen Y T; Guze A O; Stockert E; Arden K C; Boyer A D; Vittas C S; Scanlan M J; Old L	Wang, Tongtong; Hopkins, Deborah; Schmidt, Cheryl; Siliva, Sandra; Houghton, Raymond; Takita, Hiroshi; Repasky, Elizabeth; Reed, Steven G.	Identification of genes differentially over-expressed in lung squamous cell carcinoma using combination of cDNA subtraction and microarray analysis	CODEN: ONCNE; ISSN: 0950-9232
AU	ANSWER 20 OF 36 MEDLINE COPYRIGHT 2003 ACS DUPLICATE 9	Guze A O; Stockert E; Arden K C; Boyer A D; Vittas C S; Scanlan M J; Old L	Chen Y T; Guze A O; Stockert E; Arden K C; Boyer A D; Vittas C S; Scanlan M J; Old L	CIT10: a new cancer-testis (CT) antigen homologous to CT7 and the Mage family, identified by representation-difference analysis.	INTERNATIONAL JOURNAL OF CANCER, (2000) Mar 1) 85 (5) 726-32.
SO					Journal code: 0042124. ISSN: 0020-7136.
TI	ANSWER 21 OF 36 CAPLUS COPYRIGHT 2003 ACS	Chen, Yao-Tseng; Guze, Al; Tsang, Solam; Stockert, Elisabeth; Jaget,	Elik; Knuth, Alexander; Old, Lloyd J.	Isolated nucleic acid molecule encoding the cancer associated antigens CT7, ROC-2 and ROC-3, expression of the cancer associated antigens and application to cancer diagnosis and immunotherapy	CODEN: PIXXDB2
IN	ANSWER 22 OF 36 BIOSIS COPYRIGHT 2003 BIOLGICAL ABSTRACTS INC.	Wallrapp, C.; Mueller-Pillasch, F.; Mcsha, A.; Mengher, C.; Geng, M.; Solinas-Toldo, S.; Lichter, P.; Frohme, M.; Hoeftel, J.; D.; Adler, G.; Gross, T. M. (1)	Solinas-Toldo, S.; Lichter, P.; Frohme, M.; Hoeftel, J.; D.; Adler, G.; Gross, T. M. (1)	Strategies for the detection of disease genes in pancreatic cancer	TI

- Lohr, J.-Matthias, Colcher, David, Hollingsworth, Michael A., Liebe,  
 Stefan. Annals of the New York Academy of Sciences, (June 30, 1999) Vol.  
 880, pp. 122-146. Annals of the New York Academy of Sciences; Cell and  
 molecular biology of pancreatic carcinoma: Recent developments in research  
 and experimental therapy. Princeton, NJ: New York Academy of Sciences 2 East 63rd Street, New York, NY,  
 10021, USA.  
 ISSN: 0077-8923. ISBN: 1-57331-219-3 (cloth), 1-57331-220-7 (paper).  
 Meeting Info.: Baltic Pancreas Meeting on Pancreatic Carcinoma from Bench  
 to Bedside. Warnecke, Germany September 17-18, 1998  
 A novel cytoplasmic protein with RNA-binding motifs is an autoantigen in  
 human hepatocellular carcinoma.  
 Yamada, Akira; Kobayashi, Terutada; Mine, Takashii; Ushijima, Kimito;  
 Nagai, Kouji; Kawano, Koichiro; Harashima, Nanae; Niya, Fumihiko;  
 Nishida, Takaishi; Itoh, Kyogo  
 Study of HLA class I restriction and the directed antigens of cytotoxic T  
 lymphocytes at the tumor sites of ovarian cancer.  
 Hameister H, Knobel W, Adler G, Gross T M  
 Expression of the highly conserved RNA binding protein RGC in  
 embryos.  
 Chen, Yao-Tseng; Guire, Alia O.; Tsang, Solam; Stockert, Elisabeth; Jagger,  
 Elke; Kunth, Alexander; Old, Lloyd J.  
 ANSWER 26 OF 36 CAPLUS COPYRIGHT 2003 ACS  
 ANSWER 25 OF 36 MEDLINE DUPPLICATE 11  
 ISSN: 0340-7004.  
 Cancer Immunotherapy (May-June, 1999) Vol. 48, No. 2-3, pp.  
 147-152.  
 Study of HLA class I restriction and the directed antigens of cytotoxic T  
 lymphocytes at the tumor sites of ovarian cancer.  
 Mueller-Pillasch F; Pohl B; Wilda M; Lachter U; Bell M; Wallrapp C;  
 Meilexer-Pillasch F; Pohl B; Wilda M; Lachter U; Bell M; Wallrapp C;  
 Chen, Yao-Tseng; Guire, Alia O.; Tsang, Solam; Stockert, Elisabeth; Jagger,  
 Elke; Kunth, Alexander; Old, Lloyd J.  
 ANSWER 26 OF 36 CAPLUS COPYRIGHT 2003 ACS  
 ANSWER 27 OF 36 CAPLUS COPYRIGHT 2003 ACS  
 Hirota, Nobuyuki; Tomioka, Yoshito  
 Estabilishment and characterization of three human ovarian clear cell  
 carcinoma cell line in nude mice  
 Kuroume Igakka Zasshi (1998), 61(12), 334-343  
 CODEN: KIZAAL, ISSN: 0368-5810  
 ANSWER 28 OF 36 CAPLUS COPYRIGHT 2003 ACS  
 Tomioka, Yoshito  
 Estabilishment and characterization of ovarian Yolk sac tumor cell line in  
 nude mice  
 Kuroume Igakka Zasshi (1998), 61(12), 334-343  
 CODEN: KIZAAL, ISSN: 0368-5810  
 ANSWER 29 OF 36 MEDLINE DUPPLICATE 12  
 Motomurra S  
 Induction of apoptosis in ovarian carcinoma cell line by  
 gondotropin-releasing hormone agonist.  
 KURUME MEDICAL JOURNAL, (1998) 45 (1) 27-32.  
 SO TI AU L4

- Journal code: 2985210R. ISSN: 0023-5679.
- L4 ANSWER 30 OF 36 MEDLINE DUPLICATE 13
- AU Müller-Pillensch F; Lachter U; Wallrapp C; Micha A; Zimmerhackl F;  
Hameister H; Varga G; Frisse H; Buchler M; Begehr H; Viila M R; Adler G;  
Gress T M  
Cloning of a gene highly overexpressed in cancer coding for a  
novel KH-domain containing protein.  
ONCOGENE, (1997 Jun 5) 14 (22) 2729-33.  
Journal code: 8711562. ISSN: 0950-9232.
- L4 ANSWER 31 OF 36 CAPLUS COPYRIGHT 2003 ACS DUPLICATE 14
- AU Kataoka, Akio; Nishida, Takashi; Hirai, Nobuyuki; Tomotaka, Yoshito;  
Sugiyama, Toru; Yakuhashi, Michiaki  
Induction of apoptosis in ovarian carcinoma cell line by glucocorticoids,  
and sex steroid hormones  
OncoLOGY Reports, (1997), 4(6), 1249-1253  
CODEN: OCRLPM, ISSN: 1021-335X
- L4 ANSWER 32 OF 36 BIOSIS COPYRIGHT 2003 BILOGICAL ABSTRACTS INC.  
AU Müller-Pillensch, Friederike (1); Lachter, Ulrike (1); Wallrapp, Christine  
(1); Micha, Anne (1); Zimmerhackl, Frank (1); Varga, G.; Press, H.;  
Buchler, M.; Viila, M. R.; Hameister, H.; Adler, G.; Gress, Thomas  
Meeting Info.: Digestive Disease Week and the 97th Annual Meeting of the  
American Gastroenterology Association Association Washington, D.C., USA May 11-14,  
1997  
Gastroenterology, (1997) Vol. 112, No. 4 SUPPL., pp. A621.  
ISSN: 0016-5085.
- L4 ANSWER 33 OF 36 SCISEARCH COPYRIGHT 2003 ISI (R)  
AU Okina H; Kataoka A (Reprint); Sugiyama T; Nishida T; Yakuhashi M  
BestabiliShment and characterization of ovarian endometrioid adenocarcinoma  
cell line in nude mice and analyses of the immunohistochemical properties  
among the original, recurred, and heterotransplanted tumors  
PublIsher: INT JOURNAL ONCOLOGY, (FEB 1997) VOL. 10, NO. 2, PP. 311-316.  
SO International Journal of Oncology, (1994 Oct) 46 (10) 1033-40.  
ISSN: 1019-6439.
- L4 ANSWER 34 OF 36 MEDLINE DUPLICATE 15  
AU Kataoka A; Nishida T; Motomura S; Higashijima T; Ueyama T; Sugiyama T;  
Yakuhashi M  
The effects of sex steroid hormones, gonadotropins, and gonadotropin  
releasing hormone agonist on the proliferation of ovarian cancer  
TI NIPPON SANKA FUJINKA GAKKAI ZASSHI. ACTA OBSTETRICA ET GYNAECOLOGICA  
SO JAPONICA, (1994 Oct) 46 (10) 1033-40.  
Cell Line (ROC-2S).
- L4 ANSWER 35 OF 36 CAPLUS COPYRIGHT 2003 ACS  
AU Akizuki, Hideaki  
An experimental study on optimal administration of cisplatin for ovarian  
cancer  
Kurume Igakkai Zasshi (1991), 54 (5), 359-66  
CODEN: KIZAAL, ISSN: 0368-5810
- L4 ANSWER 36 OF 36 MEDLINE DUPLICATE 16  
AU Kataoka A; Yokota D; Sugiyama T; oda T; Nishida T; Yakuhashi M  
Effects of anticancer agents on 7, 12-dimethylbenz(a)anthracene induced  
ovarian cancer cell line (DMBA-OC-1) and human ovarian  
serous adenocarcinoma cell line (ROC-1S).

FILE HOME, ENTERED AT 15:48:12 ON 24 APR 2003  
 (FILE HOME, ENTERED AT 15:48:12 ON 24 APR 2003)  
 => d his  
 => s (immune (W) response or t (W) cell) and 14  
 3 files searched...  
 8 (IMMUNE (W) RESPONSE OR T (W) CELL) AND 14  
 2003 S KOC  
 0 S MUELLER-PALLASCH AND 11  
 71 S LI AND CANCER  
 36 DUP REM 13 (35 DUPLICATES REMOVED)  
 8 S (IMMUNE (W) RESPONSE OR T (W) CELL) AND 14  
 L4  
 L5  
 Wang T; Fan L; Wanabae Y; McNell P; Moulton G; Bangur C; Fangher G;  
 AU TI SO  
 Okada M; Inoue Y; Persling D; Reed S G  
 523, an RNA-binding protein as a potential therapeutic target for lung cancer.  
 BRITISH JOURNAL OF CANCER, (2003 Mar 24) 88 (6) 887-94.  
 journal code: 0370635. ISSN: 0007-0920.  
 AB  
 APPROACHES TO VACCINE-BASED IMMUNOTHERAPY OF HUMAN CANCER MAY ULTIMATELY REQUIRE TARGETS THAT ARE BOTH TUMOR-SPECIFIC AND IMMUNOGENIC.  
 IN ORDER TO GENERATE SPECIFIC ANTI-TUMOR IMMUNE RESPONSES TO LUNG CANCER, WE HAVE DEVELOPED A HIGH-PRECISEMENT LUNG CANCER-SPECIFIC PROTEINS THAT CAN BE TARGETED FOR ADJUVANT VACCINE THERAPY. BY USING A COMBINATON OF DNA SUBTRACTION AND MICROARRAY ANALYSIS, WE PREVIOUSLY REPORTED THE IDENTIFICATION OF AN RNA-BINDING PROTEIN WITHIN THE KOC FAMILY, L523, TO BE OVEREXRESSED IN SQUAMOUS CELLS OF THE LUNG. WE SHOW HERE THAT L523 IS AN ONCOFETAL PROTEIN, L523 IS NORMALLY EXPRESSED IN CANCER. AS AN ONCOFETAL PROTEIN, L523 IS IMMUNOTHERAPY OF LUNG EXHIBITS SIGNIFICANT POTENTIAL FOR VACCINE IMMUNOTHERAPY OF LUNG EARLY EMBRYONIC TISSUES, YET IT IS NOT EXPRESSED IN A HIGH-PRECISEMENT LUNG CANCER. THE SPECIFICITY OF L523 EXPRESSED IN CANCER WAS DEMONSTRATED BY BOTH mRNA AND PROTEIN MEASUREMENTS FURTHERMORE, WE SHOW THAT IMMUNOLOGICAL TOLERANCE OF L523 IS NATURALLY BROKEN IN LUNG CANCER PATIENTS, AS EVIDENCED BY DETECTABLE PLUERAL EFFUSIONS FROM LUNG CANCER PATIENTS. COLLECTIVELY, OUR STUDIES SUGGEST THAT L523 MAY BE AN IMPORTANT MARKER OF MALIGNANT PROGRESSION IN HUMAN LUNG CANCER, AND FURTHER SUGGEST THAT TREATMENT APPROACHES BASED ON L523 AS AN IMMUNOGENIC TARGET ARE WORTHY OF PURSUIT. BRITISH JOURNAL OF CANCER (2003) 88, 887-894.

L5 ANSWER 1 OF 8 MEDLINE  
 AU TI SO  
 Wang T; Fan L; Wanabae Y; McNell P; Moulton G; Bangur C; Fangher G;  
 TI SO  
 Okada M; Inoue Y; Persling D; Reed S G  
 523, an RNA-binding protein as a potential therapeutic target for lung cancer.  
 BRITISH JOURNAL OF CANCER, (2003 Mar 24) 88 (6) 887-94.  
 journal code: 0370635. ISSN: 0007-0920.  
 AB  
 APPROACHES TO VACCINE-BASED IMMUNOTHERAPY OF HUMAN CANCER MAY ULTIMATELY REQUIRE TARGETS THAT ARE BOTH TUMOR-SPECIFIC AND IMMUNOGENIC.  
 IN ORDER TO GENERATE SPECIFIC ANTI-TUMOR IMMUNE RESPONSES TO LUNG CANCER, WE HAVE DEVELOPED A HIGH-PRECISEMENT LUNG CANCER-SPECIFIC PROTEINS THAT CAN BE TARGETED FOR ADJUVANT VACCINE THERAPY. BY USING A COMBINATON OF DNA SUBTRACTION AND MICROARRAY ANALYSIS, WE PREVIOUSLY REPORTED THE IDENTIFICATION OF AN RNA-BINDING PROTEIN WITHIN THE KOC FAMILY, L523, TO BE OVEREXRESSED IN SQUAMOUS CELLS OF THE LUNG. WE SHOW HERE THAT L523 IS AN ONCOFETAL PROTEIN, L523 IS NORMALLY EXPRESSED IN CANCER. AS AN ONCOFETAL PROTEIN, L523 IS IMMUNOTHERAPY OF LUNG EXHIBITS SIGNIFICANT POTENTIAL FOR VACCINE IMMUNOTHERAPY OF LUNG EARLY EMBRYONIC TISSUES, YET IT IS NOT EXPRESSED IN A HIGH-PRECISEMENT LUNG CANCER. THE SPECIFICITY OF L523 EXPRESSED IN CANCER WAS DEMONSTRATED BY BOTH mRNA AND PROTEIN MEASUREMENTS FURTHERMORE, WE SHOW THAT IMMUNOLOGICAL TOLERANCE OF L523 IS NATURALLY BROKEN IN LUNG CANCER PATIENTS, AS EVIDENCED BY DETECTABLE PLUERAL EFFUSIONS FROM LUNG CANCER PATIENTS. COLLECTIVELY, OUR STUDIES SUGGEST THAT L523 MAY BE AN IMPORTANT MARKER OF MALIGNANT PROGRESSION IN HUMAN LUNG CANCER, AND FURTHER SUGGEST THAT TREATMENT APPROACHES BASED ON L523 AS AN IMMUNOGENIC TARGET ARE WORTHY OF PURSUIT. BRITISH JOURNAL OF CANCER (2003) 88, 887-894.

L5 ANSWER 2 OF 8 MEDLINE  
 AU TI SO  
 Zhang J Y; Chan E K; Peng X; Lu M; Wang X; Mueller F; Tan E M  
 CLINICAL IMMUNOLOGY, (2001 Aug) 100 (2) 149-56.  
 journal code: 10083537. ISSN: 1521-6616.  
 Two tumor-associated antigens, p62 and Koc, are insulin-like growth factor II (IGF-II) messenger RNA binding proteins. Autoantibodies to p62 have been detected in cancer sera but have not been detected in normal individuals. Autoantibodies to p62 have been detected in cancer sera but have not been detected in normal individuals.

reported for Koc. This study determined the extent and frequency of autoantibodies to p62 and Koc in diverse malignancies, the epitopes on the antigens, and the presence or absence of cross-reactive antibodies. Recombinant polypeptides were expressed from full-length and partial cDNA constructs and used as antigens in Western blotting, enzyme-linked immunosorbent assay, and immunoprecipitation. After identification of the epitopes, cross-absorption with recombinant polypeptides was used to determine specificity. sera from 777 patients with 10 different malignancies were analyzed. Autoantibodies to p62 were found in 11.6% and to Koc in 12.2% and cumulatively to both antigens in 20.5%, with significant difference from the control population ( $P < 0.01$ ). The immunodominant epitopes were at the amino termini of both antigens and absorbed studies showed that the majority of autoantibodies were not cross-reactive. Autoantibodies to p62 and Koc were present in approximately similar frequency in a variety of malignancies and the immune responses appeared to be independent of each other. The immune responses might be related to overexpression or dysregulation of p62 and Koc in some tumors.

L5	ANSWER 3 OF 8	MEDLINE	Guire A O; Stocker E; Arden K C; Boyer A D; Vinars C S; Scanlan M J; Old L J; Chen Y T	T1	CT10: a new cancer-testis (CT) antigen homologous to CTC and the MAGE family, identified by representational difference analysis.	SO	INTERNATIONAL JOURNAL OF CANCER, (2000 Mar 1) 85 (5) 726-32.	AB
					Journals received by [REDACTED] as [REDACTED] based research		Journal code: 0042124. ISSN: 0020-7136.	

of tumor antigens to identify potential targets for immunotherapy have led to the discovery of a significant number of immunogenic gene products, including cancer-testis (CT) antigens predominantly expressed in cancer cells and male germ cells. The search for cancer-specific antigens has been extended via the technique of peptide presentation-differential analysis and SK-MEL-37, a melanoma cell line expressing a broad range of CT antigens. Using this approach, we have isolated CT antigen genes, genes over-expressed in cancer, e.g., PRAME and KOC, and genes encoding neuroectodermal markers. The identification of CT antigen genes includes the previously defined MAGE-A6, MAGE-A4, MAGE-C1, as well as a novel gene designated MAGE-C10, which shows strong homology to CT7/MAGE-C1 both at cDNA and at genomic levels. Chromosome mapping localized CT10 to Xq27, in close proximity to CT7/MAGE-C1 and MAGE-A genes. CT10 mRNA is expressed in testes and in 20 to 30% of various human cancers. A serological survey to 30% of melanoma patients with anti-CT10 antibody, demonstrating the immunogenicity of CT10 in humans.

Zhang J Y; Chan E K; Peng X X; Tan E M  
 A novel cytoplasmic protein with RNA-binding motifs is an autoantigen in human hepatocellular carcinoma.  
 JOURNAL OF EXPERIMENTAL MEDICINE, (1999 Apr 5) 189 (7) 1101-10.  
 ISSN: 0022-1107.  
 journal code: 2985109R.  
 In hepatocellular carcinoma (HCC), autoantibodies to intracellular antigens are detected in 30-40% of patients. Patients with chronic hepatitis or liver cirrhosis develop HCC, and when this occurs, some patients excrete autoantibodies of new specificities. It has been suggested that these novel antibody responses may be immune system reactions to proteins involved in transformation-associated cellular events. One HCC serum shown to contain antibodies to unidentified cellular antigens was used to immunoscreen a cDNA expression library, and a full length cDNA clone was isolated with an open reading frame encoding 556 amino acids with a predicted molecular mass of 62 kD. The 62-kD protein was used to immunoscreen a cDNA library, and a full length cDNA clone was isolated with an open reading frame encoding 556 amino acids with a predicted molecular mass of 62 kD.





IN Anti-igenes and application to cancer diagnostics and immunotherapy  
 IN  
 PA Elike; Knuth; Alexander; Old; Lloyd J.  
 PA Ludwig Institute for Cancer Research, USA  
 SO PCT Int. Appl., 45 pp.  
 DT Patent PIXXID2  
 LA English  
 FAN. CNT 1  
 PATENT NO. KIND DATE  
 APPPLICATION NO. DATE  
 PI MO 9954738 A1 19991028 MO 1999-US5766 19990316  
 W: AU, CA, CN, JP, KR, NZ, ZA  
 RW: AT, BE, CH, CY, DE, DK, ES, FR, GB, IE, IT, LU, MC, NL,  
 US 6297364 B1 20011002 US 1998-61709 19980417  
 CA 2324975 AA 19991028 CA 1999-2324975 19990316  
 AU 9930079 A1 19991108 AU 1999-30079 19990316  
 AU 745259 B2 20020314 EP 1999-911439 19990316  
 EP 10711957 A1 20010131 EP 1999-911439 19990316  
 R: AT, BE, CH, DE, DK, ES, FR, GB, IT, LI, LU, NL, SE, MC, PT,  
 US 2002512049 T2 20020423 JP 2000-545030 19990316  
 PRAI US 1998-61709 A1 20020815 US 2001-899651 20010706  
 MO 1999-US5766 W 19990316  
 RE. CNT 5 THERE ARE 5 CITED REFERENCES AVAILABLE FOR THIS RECORD

&lt;=

GenCore version 5.1.4-p5-4578  
copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2003, 16:11:07 ; Search time: 38 Seconds  
(without alignments)  
3139.507 Million cell updates/sec

Title: US-09-897-778-176  
Perfect score: 2956  
Sequence: 1 MNKLYIGNLSSENAAPSDLES.....VQHQOQQKALQSQQPPSRRK 579

Scoring table: BloSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_21:  
 1: sp\_archaea:  
 2: sp\_bacteria:  
 3: sp\_fungi:  
 4: sp\_human:  
 5: sp\_invertebrate:  
 6: sp\_mammal:  
 7: sp\_mythic:  
 8: sp\_organelle:  
 9: sp\_phage:  
 10: sp\_plant:  
 11: sp\_rabbit:  
 12: sp\_virus:  
 13: sp\_vertebrate:  
 14: sp\_unclassified:  
 15: sp\_virus:  
 16: sp\_bacteriab:  
 17: sp\_archeap:  
 18: sp\_bacteria:  
 19: sp\_fungi:  
 20: sp\_human:  
 21: sp\_invertebrate:  
 22: sp\_mammal:  
 23: sp\_mythic:  
 24: sp\_organelle:  
 25: sp\_phage:  
 26: sp\_plant:  
 27: sp\_rabbit:  
 28: sp\_virus:  
 29: sp\_vertebrate:  
 30: sp\_unclassified:  
 31: sp\_virus:  
 32: sp\_bacteriab:  
 33: sp\_archeap:  
 34: sp\_bacteria:  
 35: sp\_fungi:  
 36: sp\_human:  
 37: sp\_invertebrate:  
 38: sp\_mammal:  
 39: sp\_mythic:  
 40: sp\_organelle:  
 41: sp\_phage:  
 42: sp\_plant:  
 43: sp\_rabbit:  
 44: sp\_virus:  
 45: sp\_vertebrate:  
 46: sp\_unclassified:  
 47: sp\_virus:  
 48: sp\_bacteriab:  
 49: sp\_archeap:  
 50: sp\_bacteria:  
 51: sp\_fungi:  
 52: sp\_human:  
 53: sp\_invertebrate:  
 54: sp\_mammal:  
 55: sp\_mythic:  
 56: sp\_organelle:  
 57: sp\_phage:  
 58: sp\_plant:  
 59: sp\_rabbit:  
 60: sp\_virus:  
 61: sp\_vertebrate:  
 62: sp\_unclassified:  
 63: sp\_virus:  
 64: sp\_bacteriab:  
 65: sp\_archeap:  
 66: sp\_bacteria:  
 67: sp\_fungi:  
 68: sp\_human:  
 69: sp\_invertebrate:  
 70: sp\_mammal:  
 71: sp\_mythic:  
 72: sp\_organelle:  
 73: sp\_phage:  
 74: sp\_plant:  
 75: sp\_rabbit:  
 76: sp\_virus:  
 77: sp\_vertebrate:  
 78: sp\_unclassified:  
 79: sp\_virus:  
 80: sp\_bacteriab:  
 81: sp\_archeap:  
 82: sp\_bacteria:  
 83: sp\_fungi:  
 84: sp\_human:  
 85: sp\_invertebrate:  
 86: sp\_mammal:  
 87: sp\_mythic:  
 88: sp\_organelle:  
 89: sp\_phage:  
 90: sp\_plant:  
 91: sp\_rabbit:  
 92: sp\_virus:  
 93: sp\_vertebrate:  
 94: sp\_unclassified:  
 95: sp\_virus:  
 96: sp\_bacteriab:  
 97: sp\_archeap:  
 98: sp\_bacteria:  
 99: sp\_fungi:  
 100: sp\_human:  
 101: sp\_invertebrate:  
 102: sp\_mammal:  
 103: sp\_mythic:  
 104: sp\_organelle:  
 105: sp\_phage:  
 106: sp\_plant:  
 107: sp\_rabbit:  
 108: sp\_virus:  
 109: sp\_vertebrate:  
 110: sp\_unclassified:  
 111: sp\_virus:  
 112: sp\_bacteriab:  
 113: sp\_archeap:  
 114: sp\_bacteria:  
 115: sp\_fungi:  
 116: sp\_human:  
 117: sp\_invertebrate:  
 118: sp\_mammal:  
 119: sp\_mythic:  
 120: sp\_organelle:  
 121: sp\_phage:  
 122: sp\_plant:  
 123: sp\_rabbit:  
 124: sp\_virus:  
 125: sp\_vertebrate:  
 126: sp\_unclassified:  
 127: sp\_virus:  
 128: sp\_bacteriab:  
 129: sp\_archeap:  
 130: sp\_bacteria:  
 131: sp\_fungi:  
 132: sp\_human:  
 133: sp\_invertebrate:  
 134: sp\_mammal:  
 135: sp\_mythic:  
 136: sp\_organelle:  
 137: sp\_phage:  
 138: sp\_plant:  
 139: sp\_rabbit:  
 140: sp\_virus:  
 141: sp\_vertebrate:  
 142: sp\_unclassified:  
 143: sp\_virus:  
 144: sp\_bacteriab:  
 145: sp\_archeap:  
 146: sp\_bacteria:  
 147: sp\_fungi:  
 148: sp\_human:  
 149: sp\_invertebrate:  
 150: sp\_mammal:  
 151: sp\_mythic:  
 152: sp\_organelle:  
 153: sp\_phage:  
 154: sp\_plant:  
 155: sp\_rabbit:  
 156: sp\_virus:  
 157: sp\_vertebrate:  
 158: sp\_unclassified:  
 159: sp\_virus:  
 160: sp\_bacteriab:  
 161: sp\_archeap:  
 162: sp\_bacteria:  
 163: sp\_fungi:  
 164: sp\_human:  
 165: sp\_invertebrate:  
 166: sp\_mammal:  
 167: sp\_mythic:  
 168: sp\_organelle:  
 169: sp\_phage:  
 170: sp\_plant:  
 171: sp\_rabbit:  
 172: sp\_virus:  
 173: sp\_vertebrate:  
 174: sp\_unclassified:  
 175: sp\_virus:  
 176: sp\_bacteriab:  
 177: sp\_archeap:  
 178: sp\_bacteria:  
 179: sp\_fungi:  
 180: sp\_human:  
 181: sp\_invertebrate:  
 182: sp\_mammal:  
 183: sp\_mythic:  
 184: sp\_organelle:  
 185: sp\_phage:  
 186: sp\_plant:  
 187: sp\_rabbit:  
 188: sp\_virus:  
 189: sp\_vertebrate:  
 190: sp\_unclassified:  
 191: sp\_virus:  
 192: sp\_bacteriab:  
 193: sp\_archeap:  
 194: sp\_bacteria:  
 195: sp\_fungi:  
 196: sp\_human:  
 197: sp\_invertebrate:  
 198: sp\_mammal:  
 199: sp\_mythic:  
 200: sp\_organelle:  
 201: sp\_phage:  
 202: sp\_plant:  
 203: sp\_rabbit:  
 204: sp\_virus:  
 205: sp\_vertebrate:  
 206: sp\_unclassified:  
 207: sp\_virus:  
 208: sp\_bacteriab:  
 209: sp\_archeap:  
 210: sp\_bacteria:  
 211: sp\_fungi:  
 212: sp\_human:  
 213: sp\_invertebrate:  
 214: sp\_mammal:  
 215: sp\_mythic:  
 216: sp\_organelle:  
 217: sp\_phage:  
 218: sp\_plant:  
 219: sp\_rabbit:  
 220: sp\_virus:  
 221: sp\_vertebrate:  
 222: sp\_unclassified:  
 223: sp\_virus:  
 224: sp\_bacteriab:  
 225: sp\_archeap:  
 226: sp\_bacteria:  
 227: sp\_fungi:  
 228: sp\_human:  
 229: sp\_invertebrate:  
 230: sp\_mammal:  
 231: sp\_mythic:  
 232: sp\_organelle:  
 233: sp\_phage:  
 234: sp\_plant:  
 235: sp\_rabbit:  
 236: sp\_virus:  
 237: sp\_vertebrate:  
 238: sp\_unclassified:  
 239: sp\_virus:  
 240: sp\_bacteriab:  
 241: sp\_archeap:  
 242: sp\_bacteria:  
 243: sp\_fungi:  
 244: sp\_human:  
 245: sp\_invertebrate:  
 246: sp\_mammal:  
 247: sp\_mythic:  
 248: sp\_organelle:  
 249: sp\_phage:  
 250: sp\_plant:  
 251: sp\_rabbit:  
 252: sp\_virus:  
 253: sp\_vertebrate:  
 254: sp\_unclassified:  
 255: sp\_virus:  
 256: sp\_bacteriab:  
 257: sp\_archeap:  
 258: sp\_bacteria:  
 259: sp\_fungi:  
 260: sp\_human:  
 261: sp\_invertebrate:  
 262: sp\_mammal:  
 263: sp\_mythic:  
 264: sp\_organelle:  
 265: sp\_phage:  
 266: sp\_plant:  
 267: sp\_rabbit:  
 268: sp\_virus:  
 269: sp\_vertebrate:  
 270: sp\_unclassified:  
 271: sp\_virus:  
 272: sp\_bacteriab:  
 273: sp\_archeap:  
 274: sp\_bacteria:  
 275: sp\_fungi:  
 276: sp\_human:  
 277: sp\_invertebrate:  
 278: sp\_mammal:  
 279: sp\_mythic:  
 280: sp\_organelle:  
 281: sp\_phage:  
 282: sp\_plant:  
 283: sp\_rabbit:  
 284: sp\_virus:  
 285: sp\_vertebrate:  
 286: sp\_unclassified:  
 287: sp\_virus:  
 288: sp\_bacteriab:  
 289: sp\_archeap:  
 290: sp\_bacteria:  
 291: sp\_fungi:  
 292: sp\_human:  
 293: sp\_invertebrate:  
 294: sp\_mammal:  
 295: sp\_mythic:  
 296: sp\_organelle:  
 297: sp\_phage:  
 298: sp\_plant:  
 299: sp\_rabbit:  
 300: sp\_virus:  
 301: sp\_vertebrate:  
 302: sp\_unclassified:  
 303: sp\_virus:  
 304: sp\_bacteriab:  
 305: sp\_archeap:  
 306: sp\_bacteria:  
 307: sp\_fungi:  
 308: sp\_human:  
 309: sp\_invertebrate:  
 310: sp\_mammal:  
 311: sp\_mythic:  
 312: sp\_organelle:  
 313: sp\_phage:  
 314: sp\_plant:  
 315: sp\_rabbit:  
 316: sp\_virus:  
 317: sp\_vertebrate:  
 318: sp\_unclassified:  
 319: sp\_virus:  
 320: sp\_bacteriab:  
 321: sp\_archeap:  
 322: sp\_bacteria:  
 323: sp\_fungi:  
 324: sp\_human:  
 325: sp\_invertebrate:  
 326: sp\_mammal:  
 327: sp\_mythic:  
 328: sp\_organelle:  
 329: sp\_phage:  
 330: sp\_plant:  
 331: sp\_rabbit:  
 332: sp\_virus:  
 333: sp\_vertebrate:  
 334: sp\_unclassified:  
 335: sp\_virus:  
 336: sp\_bacteriab:  
 337: sp\_archeap:  
 338: sp\_bacteria:  
 339: sp\_fungi:  
 340: sp\_human:  
 341: sp\_invertebrate:  
 342: sp\_mammal:  
 343: sp\_mythic:  
 344: sp\_organelle:  
 345: sp\_phage:  
 346: sp\_plant:  
 347: sp\_rabbit:  
 348: sp\_virus:  
 349: sp\_vertebrate:  
 350: sp\_unclassified:  
 351: sp\_virus:  
 352: sp\_bacteriab:  
 353: sp\_archeap:  
 354: sp\_bacteria:  
 355: sp\_fungi:  
 356: sp\_human:  
 357: sp\_invertebrate:  
 358: sp\_mammal:  
 359: sp\_mythic:  
 360: sp\_organelle:  
 361: sp\_phage:  
 362: sp\_plant:  
 363: sp\_rabbit:  
 364: sp\_virus:  
 365: sp\_vertebrate:  
 366: sp\_unclassified:  
 367: sp\_virus:  
 368: sp\_bacteriab:  
 369: sp\_archeap:  
 370: sp\_bacteria:  
 371: sp\_fungi:  
 372: sp\_human:  
 373: sp\_invertebrate:  
 374: sp\_mammal:  
 375: sp\_mythic:  
 376: sp\_organelle:  
 377: sp\_phage:  
 378: sp\_plant:  
 379: sp\_rabbit:  
 380: sp\_virus:  
 381: sp\_vertebrate:  
 382: sp\_unclassified:  
 383: sp\_virus:  
 384: sp\_bacteriab:  
 385: sp\_archeap:  
 386: sp\_bacteria:  
 387: sp\_fungi:  
 388: sp\_human:  
 389: sp\_invertebrate:  
 390: sp\_mammal:  
 391: sp\_mythic:  
 392: sp\_organelle:  
 393: sp\_phage:  
 394: sp\_plant:  
 395: sp\_rabbit:  
 396: sp\_virus:  
 397: sp\_vertebrate:  
 398: sp\_unclassified:  
 399: sp\_virus:  
 400: sp\_bacteriab:  
 401: sp\_archeap:  
 402: sp\_bacteria:  
 403: sp\_fungi:  
 404: sp\_human:  
 405: sp\_invertebrate:  
 406: sp\_mammal:  
 407: sp\_mythic:  
 408: sp\_organelle:  
 409: sp\_phage:  
 410: sp\_plant:  
 411: sp\_rabbit:  
 412: sp\_virus:  
 413: sp\_vertebrate:  
 414: sp\_unclassified:  
 415: sp\_virus:  
 416: sp\_bacteriab:  
 417: sp\_archeap:  
 418: sp\_bacteria:  
 419: sp\_fungi:  
 420: sp\_human:  
 421: sp\_invertebrate:  
 422: sp\_mammal:  
 423: sp\_mythic:  
 424: sp\_organelle:  
 425: sp\_phage:  
 426: sp\_plant:  
 427: sp\_rabbit:  
 428: sp\_virus:  
 429: sp\_vertebrate:  
 430: sp\_unclassified:  
 431: sp\_virus:  
 432: sp\_bacteriab:  
 433: sp\_archeap:  
 434: sp\_bacteria:  
 435: sp\_fungi:  
 436: sp\_human:  
 437: sp\_invertebrate:  
 438: sp\_mammal:  
 439: sp\_mythic:  
 440: sp\_organelle:  
 441: sp\_phage:  
 442: sp\_plant:  
 443: sp\_rabbit:  
 444: sp\_virus:  
 445: sp\_vertebrate:  
 446: sp\_unclassified:  
 447: sp\_virus:  
 448: sp\_bacteriab:  
 449: sp\_archeap:  
 450: sp\_bacteria:  
 451: sp\_fungi:  
 452: sp\_human:  
 453: sp\_invertebrate:  
 454: sp\_mammal:  
 455: sp\_mythic:  
 456: sp\_organelle:  
 457: sp\_phage:  
 458: sp\_plant:  
 459: sp\_rabbit:  
 460: sp\_virus:  
 461: sp\_vertebrate:  
 462: sp\_unclassified:  
 463: sp\_virus:  
 464: sp\_bacteriab:  
 465: sp\_archeap:  
 466: sp\_bacteria:  
 467: sp\_fungi:  
 468: sp\_human:  
 469: sp\_invertebrate:  
 470: sp\_mammal:  
 471: sp\_mythic:  
 472: sp\_organelle:  
 473: sp\_phage:  
 474: sp\_plant:  
 475: sp\_rabbit:  
 476: sp\_virus:  
 477: sp\_vertebrate:  
 478: sp\_unclassified:  
 479: sp\_virus:  
 480: sp\_bacteriab:  
 481: sp\_archeap:  
 482: sp\_bacteria:  
 483: sp\_fungi:  
 484: sp\_human:  
 485: sp\_invertebrate:  
 486: sp\_mammal:  
 487: sp\_mythic:  
 488: sp\_organelle:  
 489: sp\_phage:  
 490: sp\_plant:  
 491: sp\_rabbit:  
 492: sp\_virus:  
 493: sp\_vertebrate:  
 494: sp\_unclassified:  
 495: sp\_virus:  
 496: sp\_bacteriab:  
 497: sp\_archeap:  
 498: sp\_bacteria:  
 499: sp\_fungi:  
 500: sp\_human:  
 501: sp\_invertebrate:  
 502: sp\_mammal:  
 503: sp\_mythic:  
 504: sp\_organelle:  
 505: sp\_phage:  
 506: sp\_plant:  
 507: sp\_rabbit:  
 508: sp\_virus:  
 509: sp\_vertebrate:  
 510: sp\_unclassified:  
 511: sp\_virus:  
 512: sp\_bacteriab:  
 513: sp\_archeap:  
 514: sp\_bacteria:  
 515: sp\_fungi:  
 516: sp\_human:  
 517: sp\_invertebrate:  
 518: sp\_mammal:  
 519: sp\_mythic:  
 520: sp\_organelle:  
 521: sp\_phage:  
 522: sp\_plant:  
 523: sp\_rabbit:  
 524: sp\_virus:  
 525: sp\_vertebrate:  
 526: sp\_unclassified:  
 527: sp\_virus:  
 528: sp\_bacteriab:  
 529: sp\_archeap:  
 530: sp\_bacteria:  
 531: sp\_fungi:  
 532: sp\_human:  
 533: sp\_invertebrate:  
 534: sp\_mammal:  
 535: sp\_mythic:  
 536: sp\_organelle:  
 537: sp\_phage:  
 538: sp\_plant:  
 539: sp\_rabbit:  
 540: sp\_virus:  
 541: sp\_vertebrate:  
 542: sp\_unclassified:  
 543: sp\_virus:  
 544: sp\_bacteriab:  
 545: sp\_archeap:  
 546: sp\_bacteria:  
 547: sp\_fungi:  
 548: sp\_human:  
 549: sp\_invertebrate:  
 550: sp\_mammal:  
 551: sp\_mythic:  
 552: sp\_organelle:  
 553: sp\_phage:  
 554: sp\_plant:  
 555: sp\_rabbit:  
 556: sp\_virus:  
 557: sp\_vertebrate:  
 558: sp\_unclassified:  
 559: sp\_virus:  
 560: sp\_bacteriab:  
 561: sp\_archeap:  
 562: sp\_bacteria:  
 563: sp\_fungi:  
 564: sp\_human:  
 565: sp\_invertebrate:  
 566: sp\_mammal:  
 567: sp\_mythic:  
 568: sp\_organelle:  
 569: sp\_phage:  
 570: sp\_plant:  
 571: sp\_rabbit:  
 572: sp\_virus:  
 573: sp\_vertebrate:  
 574: sp\_unclassified:  
 575: sp\_virus:  
 576: sp\_bacteriab:  
 577: sp\_archeap:  
 578: sp\_bacteria:  
 579: sp\_fungi:  
 580: sp\_human:  
 581: sp\_invertebrate:  
 582: sp\_mammal:  
 583: sp\_mythic:  
 584: sp\_organelle:  
 585: sp\_phage:  
 586: sp\_plant:  
 587: sp\_rabbit:  
 588: sp\_virus:  
 589: sp\_vertebrate:  
 590: sp\_unclassified:  
 591: sp\_virus:  
 592: sp\_bacteriab:  
 593: sp\_archeap:  
 594: sp\_bacteria:  
 595: sp\_fungi:  
 596: sp\_human:  
 597: sp\_invertebrate:  
 598: sp\_mammal:  
 599: sp\_mythic:  
 600: sp\_organelle:  
 601: sp\_phage:  
 602: sp\_plant:  
 603: sp\_rabbit:  
 604: sp\_virus:  
 605: sp\_vertebrate:  
 606: sp\_unclassified:  
 607: sp\_virus:  
 608: sp\_bacteriab:  
 609: sp\_archeap:  
 610: sp\_bacteria:  
 611: sp\_fungi:  
 612: sp\_human:  
 613: sp\_invertebrate:  
 614: sp\_mammal:  
 615: sp\_mythic:  
 616: sp\_organelle:  
 617: sp\_phage:  
 618: sp\_plant:  
 619: sp\_rabbit:  
 620: sp\_virus:  
 621: sp\_vertebrate:  
 622: sp\_unclassified:  
 623: sp\_virus:  
 624: sp\_bacteriab:  
 625: sp\_archeap:  
 626: sp\_bacteria:  
 627: sp\_fungi:  
 628: sp\_human:  
 629: sp\_invertebrate:  
 630: sp\_mammal:  
 631: sp\_mythic:  
 632: sp\_organelle:  
 633: sp\_phage:  
 634: sp\_plant:  
 635: sp\_rabbit:  
 636: sp\_virus:  
 637: sp\_vertebrate:  
 638: sp\_unclassified:  
 639: sp\_virus:  
 640: sp\_bacteriab:  
 641: sp\_archeap:  
 642: sp\_bacteria:  
 643: sp\_fungi:  
 644: sp\_human:  
 645: sp\_invertebrate:  
 646: sp\_mammal:  
 647: sp\_mythic:  
 648: sp\_organelle:  
 649: sp\_phage:  
 650: sp\_plant:  
 651: sp\_rabbit:  
 652: sp\_virus:  
 653: sp\_vertebrate:  
 654: sp\_unclassified:  
 655: sp\_virus:  
 656: sp\_bacteriab:  
 657: sp\_archeap:  
 658: sp\_bacteria:  
 659: sp\_fungi:  
 660: sp\_human:  
 661: sp\_invertebrate:  
 662: sp\_mammal:  
 663: sp\_mythic:  
 664: sp\_organelle:  
 665: sp\_phage:  
 666: sp\_plant:  
 667: sp\_rabbit:  
 668: sp\_virus:  
 669: sp\_vertebrate:  
 670: sp\_unclassified:  
 671: sp\_virus:  
 672: sp\_bacteriab:  
 673: sp\_archeap:  
 674: sp\_bacteria:  
 675: sp\_fungi:  
 676: sp\_human:  
 677: sp\_invertebrate:  
 678: sp\_mammal:  
 679: sp\_mythic:  
 680: sp\_organelle:  
 681: sp\_phage:  
 682: sp\_plant:  
 683: sp\_rabbit:  
 684: sp\_virus:  
 685: sp\_vertebrate:  
 686: sp\_unclassified:  
 687: sp\_virus:  
 688: sp\_bacteriab:  
 689: sp\_archeap:  
 690: sp\_bacteria:  
 691: sp\_fungi:  
 692: sp\_human:  
 693: sp\_invertebrate:  
 694: sp\_mammal:  
 695: sp\_mythic:  
 696: sp\_organelle:  
 697: sp\_phage:  
 698: sp\_plant:  
 699: sp\_rabbit:  
 700: sp\_virus:  
 701: sp\_vertebrate:  
 702: sp\_unclassified:  
 703: sp\_virus:  
 704: sp\_bacteriab:  
 705: sp\_archeap:  
 706: sp\_bacteria:  
 707: sp\_fungi:  
 708: sp\_human:  
 709: sp\_invertebrate:  
 710: sp\_mammal:  
 711: sp\_mythic:  
 712: sp\_organelle:  
 713: sp\_phage:  
 714: sp\_plant:  
 715: sp\_rabbit:  
 716: sp\_virus:  
 717: sp\_vertebrate:  
 718: sp\_unclassified:  
 719: sp\_virus:  
 720: sp\_bacteriab:  
 721: sp\_archeap:  
 722: sp\_bacteria:  
 723: sp\_fungi:  
 724: sp\_human:  
 725: sp\_invertebrate:  
 726: sp\_mammal:  
 727: sp\_mythic:  
 728: sp\_organelle:  
 729: sp\_phage:  
 730: sp\_plant:  
 731: sp\_rabbit:  
 732: sp\_virus:  
 733: sp\_vertebrate:  
 734: sp\_unclassified:  
 735: sp\_virus:  
 736: sp\_bacteriab:  
 737: sp\_archeap:  
 738: sp\_bacteria:  
 739: sp\_fungi:  
 740: sp\_human:  
 741: sp\_invertebrate:  
 742: sp\_mammal:  
 743: sp\_mythic:  
 744: sp\_organelle:  
 745: sp\_phage:  
 746: sp\_plant:  
 747: sp\_rabbit:  
 748: sp\_virus:  
 749: sp\_vertebrate:  
 750: sp\_unclassified:  
 751: sp\_virus:  
 752: sp\_bacteriab:  
 753: sp\_archeap:  
 754: sp\_bacteria:  
 755: sp\_fungi:  
 756: sp\_human:  
 757: sp\_invertebrate:  
 758: sp\_mammal:  
 759: sp\_mythic:  
 760: sp\_organelle:  
 761: sp\_phage:  
 762: sp\_plant:  
 763: sp\_rabbit:  
 764: sp\_virus:  
 765: sp\_vertebrate:  
 766: sp\_unclassified:  
 767: sp\_virus:  
 768: sp\_bacteriab:  
 769: sp\_archeap:  
 770: sp\_bacteria:  
 771: sp\_fungi:  
 772: sp\_human:  
 773: sp\_invertebrate:  
 774: sp\_mammal:  
 775: sp\_mythic:  
 776: sp\_organelle:  
 777: sp\_phage:  
 778: sp\_plant:  
 779: sp\_rabbit:  
 780: sp\_virus:  
 781: sp\_vertebrate:  
 782: sp\_unclassified:  
 783: sp\_virus:  
 784: sp\_bacteriab:  
 785: sp\_archeap:  
 786: sp\_bacteria:  
 787: sp\_fungi:  
 788: sp\_human:  
 789: sp\_invertebrate:  
 790: sp\_mammal:  
 791: sp\_mythic:  
 792: sp\_organelle:  
 793: sp\_phage:  
 794: sp\_plant:  
 795: sp\_rabbit:  
 796: sp\_virus:  
 797: sp\_vertebrate:  
 798: sp\_unclassified:  
 799: sp\_virus:  
 800: sp\_bacteriab:  
 801: sp\_archeap:  
 802: sp\_bacteria:  
 803: sp\_fungi:  
 804: sp\_human:  
 805: sp\_invertebrate:  
 806: sp\_mammal:  
 807: sp\_mythic:  
 808: sp\_organelle:  
 809: sp\_phage:  
 810: sp\_plant:  
 811: sp\_rabbit:  
 812: sp\_virus:  
 813: sp\_vertebrate:  
 814: sp\_unclassified:  
 815: sp\_virus:  
 816: sp\_bacteriab:  
 817: sp\_archeap:  
 818: sp\_bacteria:  
 819: sp\_fungi:  
 820: sp\_human:  
 821: sp\_invertebrate:  
 822: sp\_mammal:  
 823: sp\_mythic:  
 824: sp\_organelle:  
 825: sp\_phage:  
 826: sp\_plant:  
 827: sp\_rabbit:  
 828: sp\_virus:  
 829: sp\_vertebrate:  
 830: sp\_unclassified:  
 831: sp\_virus:  
 832: sp\_bacteriab:  
 833: sp\_archeap:  
 834: sp\_bacteria:  
 835: sp\_fungi:  
 836: sp\_human:  
 837: sp\_invertebrate:  
 838: sp\_mammal:  
 839: sp\_mythic:  
 840: sp\_organelle:  
 841: sp\_phage:  
 842: sp\_plant:  
 843: sp\_rabbit:  
 844: sp\_virus:  
 845: sp\_vertebrate:  
 846: sp\_unclassified:  
 847: sp\_virus:  
 848: sp\_bacteriab:  
 849: sp\_archeap:  
 850: sp\_bacteria:  
 851: sp\_fungi:  
 852: sp\_human:  
 853: sp\_invertebrate:  
 854: sp\_mammal:  
 855: sp\_mythic:  
 856: sp\_organelle:  
 857: sp\_phage:  
 858: sp\_plant:  
 859: sp\_rabbit:  
 860: sp\_virus:  
 861: sp\_vertebrate:  
 862: sp\_unclassified:  
 863: sp\_virus:  
 864: sp\_bacteriab:  
 865: sp\_archeap:  
 866: sp\_bacteria:  
 867: sp\_fungi:  
 868: sp\_human:  
 869: sp\_invertebrate:  
 870: sp\_mammal:  
 871: sp\_mythic:  
 872: sp\_organelle:  
 873: sp\_phage:  
 874: sp\_plant:  
 875: sp\_rabbit:  
 876: sp\_virus:  
 877: sp\_vertebrate:  
 878: sp\_unclassified:  
 879: sp\_virus:  
 880: sp\_bacteriab:  
 881: sp\_archeap:  
 882: sp\_bacteria:  
 883: sp\_fungi:  
 884: sp\_human:  
 885: sp\_invertebrate:  
 886: sp\_mammal:  
 887: sp\_mythic:  
 888: sp\_organelle:  
 889: sp\_phage:  
 890: sp\_plant:  
 891: sp\_rabbit

1	MNKLYIGNLSNAAPSDESTFKDAKIPVSQPFLYRKTYAFVYDPCPDESWALKAIABLSGK	60
1	MNKLYIGNLSNAAPSDESTFENDAKIPVSQPFLYRKTYAFVYDPCPDESWALKAIABLSGK	60
61	IELGKPIEVHSVKQRTRKLQIQRNTPPHILOWEVLDSILYGGVYESCEQVNTDSETA	120
61	IELHGKPIEVHSVKQRTRKLQIQRNTPPHILOWEVLDSILYGGVYESCEQVNTDSETA	120
61	VNVYNTSSKDQARQALDKLNGFQLENFTLKVAIYIDEMAAQNPQLQPRGRGLQGRGSS	180
121	VNVYNTSSKDQARQALDKLNGFQLENFTLKVAIYIDEMAAQNPQLQPRGRGLQGRGSS	180
121	VNVYNTSSKDQARQALDKLNGFQLENFTLKVAIYIDEMAAQNPQLQPRGRGLQGRGSS	180
181	RQGPSSVSKQPCDPLRULVPTQFYGAIIKGEGATINTQTQSKIDVRENAGAA	240
181	RQGPSSVSKQPCDPLRULVPTQFYGAIIKGEGATINTQTQSKIDVRENAGAA	240
241	EKSITLSPBTSACKSILEIMHKEARODIKFTETIPKILAHNNFVGRLIGKEGRNLK	300
241	EKSITLSPBTSACKSILEIMHKEARODIKFTETIPKILAHNNFVGRLIGKEGRNLK	300
301	KIEQDTDKITISPLQELTYNERTIVKGNEVCAKEEIMKKIRESYENDIASMNL	360
301	KIEQDTDKITISPLQELTYNERTIVKGNEVCAKEEIMKKIRESYENDIASMNL	360
361	QAHJPGLNLNALGLEPPSGMPPTSGPSSAMTTPYQFQESETETVHQFIPALSGVAI	420
361	QAHJPGLNLNALGLPPSGMPPTSGPSSAMTTPYQFQESETETVHQFIPALSGVAI	420
421	IGKOGHIIKOLSRFGASIKIAPEADPKYRMVITGPPEAOFKAGRTYKIKEENF	480
421	IGKOGHIIKOLSRFGASIKIAPEADPKYRMVITGPPEAOFKAGRTYKIKEENF	480
481	SPKEEVKLEAHTRVSEAGRVIGKGSKTVNELQNUSSAEVYPRDQTPDNDQVVKIT	540
481	SPKEEVKLEAHTRVSEAGRVIGKGSKTVNELQNUSSAEVYPRDQTPDNDQVVKIT	540
541	GHFYACQAQRKIQELTQVKQROOQKALQSSPPQRK	579
541	GHFYACQAQRKIQELTQVKQROOQKALQSSPPQRK	579

RESULTS 2

CPN#	PRELIMINARY;	PRT;	579 AA.
Q9CPNB			
Q9CPB8;			
01-JUN-2001	(TRIMBLRL. 17, Created)		
01-JUN-2001	(TRIMBLRL. 17, Last sequence update)		
01-JUN-2002	(TRIMBLRL. 21, Last annotation update)		
10 days embryo cDNA, RIKEN full length enriched library, clones:2610036B18, full insert sequence (19f2 mRNA-binding protein 3).			
IGF2BP3 OR 2610101N1Rik OR MIM3.			
Mus musculus (Mouse)			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
NCBI TaxID=10090;			
[1]			
SEQUENCE FROM N A.			
STRANUM=57BL/6J; ISSUE=EMBRYO;			
MEDLINE=1083660; PubMed=1127851;			
Kawa J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
Akawa T., Hara A., Fukunaga Y., Konno H., Adachi J., Fukuda S.,			
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamamoto I.,			
Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Salo R.,			
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,			
Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quachembush J.,			
Schriml L. M., Staubli F., Suzuki I., Tonita N., Wagner L., Washio T.,			
Sakai Y., Okido T., Furuno M., Kono H., Baijareli R., Barsh G.,			
Blake J., Botfield D., Bojunga N., Carninci P., De Bonaldo M.F.,			
Brownstein M.J., Bult C., Fletcher C., Fujita M., Garibaldi M.,			
Brownstein S., Hill D., Hoffmann M., Hume D.A., Karimova M., Lee N H			

RESULT 3 057526	SEQUENCE FROM N.A. RX MEDLINE=92249652; PubMed=1577195;	PRT; 593 AA.	
ID 057526;	PRELIMINARY;	PRT;	593 AA.
AC OS7526;			
DT 01-JUN-1998	(TREMBLrel. 06; Created)		
DT 01-JUN-1998	(TREMBLrel. 06; Last sequence update)		
DT 01-JUN-2002	(TREMBLrel. 21; Last annotation update)		
DR KH domain containing transcription factor B3.			
OS Xenopus laevis (African clawed frog).			
OC Bokaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Amphibia; Batracchia; Craniata; Xenopus.			
OC Xenopodinae; Xenopus.			
OX NCBI_TAXID=8355;			
RN [1]			
RP	SEQUENCE FROM N.A.		
RA Peaff S.L.; Taylor W.L.;			
RT "characterization of a Xenopus oocyte factor that binds to a developmentally regulated cis-element in the TFLII gene.";			
RL Dev. Biol. 151:306-316(1992).			
RN [2]			
RP	SEQUENCE FROM N.A.		
RA Griffin D.; Taylor W.L.;			
RL Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.			
RN [3]			
RP	SEQUENCE FROM N.A.		
RA Havin L., Git A., Elisha Z., Oberman F., Yaniv K.,			
RA Pressman Schwartz S., Standart N.M., Yisraeli J.K.;			
RL Genes Dev. 0:0-01998;			
DR AF042353; AB9757.1; -;			
EMBL; AF042353; AB9757.1; -;			
DR InterPro; IPR004087; KH_dom.			
DR InterPro; IPR004087; KH_type_1.			
DR Pfam; PF00013; KH_domain; 4.			
DR SMART; SM00322; KH; 4.			
DR SMART; SM00360; RRM; 2.			
DR PROSITE; PS50084; KH_TYPE_1; 4.			
DR PROSITE; PS50102; RRM; 2.			
DR PROSITE; PS50102; RRM; 2.			
DR PROSITE; PS50102; RRM; 2.			
SQ 593 AA; 65385 MW;			
SEQUENCE 593 AA; 65385 MW;			
Query Match 84.0% Score 2483; DB 13; Length 593; Best Local Similarity 82.1%; Pred. No. 3.6e-156; Gaps 6; Matches 487; Conservative 47; Mismatches 45; Indels 14; Gaps 6;			
QY 1 MKLYTGNSENAPSDLESIPIKDKAPIPVSPFLVKTGYAFVDCPDESALKALPSK 60			
1 MKLYTGNSENAPSDLESIPIKDKAPIPVSPFLVKTGYAFVDCPDESALKALPSK 60			
Db 1 MKLYTGNSENAPSDLESIPIKDKAPIPVSPFLVKTGYAFVDCPDESALKALPSK 60			
QY 61 VELHGPTEVEHSVKQRIRKLQINMPHLQEVLLSLLVQGVYSECVNTDSEA 120.			
Db 61 VELHGPTEVEHSVKQRIRKLQINMPHLQEVLLSLLVQGVYSECVNTDSEA 120.			
QY 121 VVNVTYSSKQDQARQLDKLNGFLENFTLVAYIPDEMAQQNPLQO-----PRGR 172			
Db 121 VVNVTYSSKQDQARQLDKLNGFLENFTLVAYIPDEMAQQNPLQO-----PRGR 172			
QY 122 HRKENGAAKSITILSRPEGTSAAACKSLRKVTPDENATPQPSQQQQPQGR 180			
Db 122 HRKENGAAKSITILSRPEGTSAAACKSLRKVTPDENATPQPSQQQQPQGR 180			
QY 123 GLGORSSRQSPSVSKQKP-CDLPRLLLVPTQFGAIGKEGATITRNTKOTOSKID 231			
Db 123 GLGORSSRQSPSVSKQKP-CDLPRLLLVPTQFGAIGKEGATITRNTKOTOSKID 231			
QY 124 HRKENGAAKSITILSRPEGTSAAACKSLRKVTPDENATPQPSQQQQPQGR 300			
Db 124 HRKENGAAKSITILSRPEGTSAAACKSLRKVTPDENATPQPSQQQQPQGR 300			
QY 125 IKGEGNLKKEIODDTKTTISPLQETLTINPERTVYGNNTCAKEEEIMKKIRESY 351			
Db 125 IKGEGNLKKEIODDTKTTISPLQETLTINPERTVYGNNTCAKEEEIMKKIRESY 351			
QY 126 IKGEGNLKKEIODDTKTTISPLQETLTINPERTVYGNNTCAKEEEIMKKIRESY 360			
Db 126 IKGEGNLKKEIODDTKTTISPLQETLTINPERTVYGNNTCAKEEEIMKKIRESY 360			
QY 127 ENDIASMLQAHLPGLNINALGLFP-TSGMPPPTSCBPSAMT-PPYPQF-EQSETETV 408			
Db 127 ENDIASMLQAHLPGLNINALGLFP-TSGMPPPTSCBPSAMT-PPYPQF-EQSETETV 408			
QY 128 RGFQRPARQSGPAAARPCKPOTEVPLMLVPTQFGAIGKEGATITRNTKOTOSKID 240			
Db 128 RGFQRPARQSGPAAARPCKPOTEVPLMLVPTQFGAIGKEGATITRNTKOTOSKID 240			

QY	231	VIRKENAGAAEKSTTILSTPEGTSAACKSILEMHEAQDIFTTEIPIKILAHNNFVGR	290	Db	121	VNVRYGARDQARMDKLNGFLMENYALRVSYIDETAAADAP- -AVGRRGFNRGPP	178
Db	241	IRKENENGAAKPITHSPPEGSAACHTIMEAQDIFTKEEIPKILAHNNFVGR	300	Oy	181	RQGSPGVSKQ PCIDLPLRLVPTQVGAIIGKEGATIRNTKOTSKIDVHRKENAGA	239
Oy	291	LICKEGRNLKKEQDNTKTTISPLQELKTLYNPERITIVKGNTCAKEEEIMKKRES	350	Db	179	RQGSPGALKPLQDSDPLRLVPTQVGAIIGKEGATIRNTKOTSKIDVHRKENAGA	238
Db	301	LICKEGRNLKKEQDNTKTTISPLQELKTLYNPERITIVKGNTCAKEEEIMKKRES	360	Oy	240	AEKSTTILSTPEGTSAACKSILEMHEAQDIFTTEIPIKILAHNNFVGR LIGKEGRNL	293
Oy	351	YENDIASMNLOAHLPGLNNAHLGLFP- PTSMPPPTSGPES- AMTPYPQF- EQSETET 407	411	Db	239	AEKPTVHSPCGOSSACRNIMOREAIDTKTEEIPIKILAHNNFVGR LIGKEGRNL	298
Db	361	YENDIASMNLOAHLPGLNNAHLGLFPSSSSMPPSVGCPSPSTSSTVPPFCQPOSET	420	Oy	300	KLEQDTDKSITISPLQELTLYNPERITIVKGNTCAKEEIMKKRES YENDIASMN	359
Oy	408	WOFIPALSGAATIKQGOHKIQLSRFGASTIKTAAEADPKYRMVITGPPEAQPKAQ	467	Db	299	KTEQDQTDTITISPLQDLYTLYNPERITIVKGTLACAAEETMKVRESYNDVAAAMH	358
Db	421	WIFHPALAVAYAISKQGOHKIQLSRFGASTIKTAAEADPKYRMVITGPPEAQPKAQ	480	Oy	360	LAHLJPGLNNAHLGLFP- PTSG- -NPPTSGPPSAMTPYPQF- EQSETETVHQFIPAL	415
Oy	468	GRIGYKIKEENFVSPEKEVKLEAHIRVPSFAAGRVIGKGKTVNELQNLSSAEEVYPRDQ	527	Db	359	LSQNLIPLGLNNAHLGLFPQAASGGISPSVSVQSGPQAQGYQSFGAOMESETVHJFIPAL	418
Db	481	GRIGYKIKEENFVSPEKEVKLEAHIRVPSFAAGRVIGKGKTVNELQNLSSAEEVYPRDQ	540	Oy	416	SVGATIGKQGQKHQLSRFGASISKIAKAPAEPAKDVARYTITCPEAQFKAQGRYJKIK	475
Oy	528	TFDENDOYVVTGTHFYACQVAQKRCIOEILTOK- -QHOQOKALQSGPQSRRK	579	Db	419	AVGATIGKQGQKHQLSRFGASISKIAKAPDGIDAQMRVYLTSPPCEAQFKQGRFGLK	478
Db	541	TFDENDEVVVVTKTGHFYASQLAQRKIQELAQYRQQQOKTQYSQFQPRRK	594	Oy	476	BENFVSPKEEVKLEAHIRVPSFAAGRIVIGKGKTVNELQNLSSAEEVYVPRDQTPDNDQV	535
Db				Db	479	EENFEGPKPEEVKLEAHIRVPSFAAGRIVIGKGKTVNELQNLSSAEEVYVPRDQTPDNDQV	538
RESULT 5							
Q9PRB0		PRELIMINARY;	PRT;	PRT;	576	AA.	
ID	Q9PRB0;				042254	PRELIMINARY;	
AC	Q9PRB0;				AC		
DT	01-MAY-2000	(TREMBrel. 13, Created)			DT	01-JAN-1998 (TREMBrel. 05, Created)	
DT	01-MAY-2000	(TREMBrel. 13, Last sequence update)			DT	01-JAN-1998 (TREMBrel. 05, Last sequence update)	
DE	01-JUN-2002	(TREMBrel. 21, Last annotation update)			DT	01-JUN-2002 (TREMBrel. 21, Last annotation update)	
GN					DE		
OS					DE		
Brachydanio rerio	(zebra danio).				ZIPBP1		
Actinopterygii	Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				GN		
Cyprinidae	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;				Gal		
RN	Danio				Archaosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;		
SEQUENCE FROM N.A.							
RA	Zhang Q., Yaniv K., Oberman F., Wolke U., Git A., Fromer M.,				NCBI_TaxID=9031;		
RA	Taylor W., Meyer D., Stoward N., Raz E., Yisraeli J.K.;				RN [1]		
RA	"g1 RBP intracellular distribution and evolutionarily conserved expression suggest multiple roles during development."				SEQUENCE FROM N.A.		
RT					MEDLINE=97220007; PubMed=9121465;		
RT					RX		
RT					Ross A.F., Oleynikov Y.S., Kisieliuskis E.H., Taneja K.L., Singer R.H.;		
RT					RR	"Characterization of a beta-catenin mRNA zipcode-binding protein."	
RT					RL	Mol. Cell. Biol. 17:258-265 (1997).	
RT					EMBL; AF026527; AB022951;		
RT					DR	IntersPro: IPR004087; KH_dom.	
RT					DR	InterPro: IPR004088; KH_type_1.	
RT					DR	InterPro: IPR000504; RNA_rec_mot.	
RT					DR	Pfam: PF00013; KH_domain; 4.	
RT					DR	Pfam: PF00076; rrm; 2.	
RT					DR	SMART: SM00322; KH; 4.	
RT					DR	PROSITE: PS50084; KH_TYPE_1; 4.	
RT					DR	PROSITE: PS50102; RRM; 2.	
RT					DR	PROSITE: PS50030; RRM; RNP_1; FALSE_NEG	
SQ	SEQUENCE 582 AA;	63351 MW;	9DAE63200681B306 CRC64;	SQ	SEQUENCE 576 AA;	63271 MW;	01AAFA2D81JC8811 CRC64;
Query Match Similarity 76.5%; Best Local Similarity 77.1%; Matches 450; Conservation 50; Mismatches 77; Indels 5; Gaps 5;							
QY	1 MNKLYIGNLSENAAPSDELESITFDKAPIPVSGPPLVTKGYAFVDCPDESWALKATEALSGK	60	Query Match Similarity 75.1%; Best Local Similarity 74.7%; Matches 437; Conservation 61; Mismatches 72; Indels 15; Gaps				
Db	1 MNKLYIGNLSENAAPSDELESITFDKAPIPVSGPPLVTKGYAFVDCPDESWALKATEALSGK	60	1 MNKLYIGNLSENAAPSDELESITFDKAPIPVSGPPLVTKGYAFVDCPDESWALKATEALSGK	60	DR PROSITE: PS50084; KH_TYPE_1; 4.		
QY	61 IELHGKPIEVHESVPKQRKIQLQRINNPHLQWEVLDSLLVQYGVVESCBQVNTSETA	120	DR PROSITE: PS50102; RRM; 2.				
Db	61 VELHGKVLEVEHESVPKQRKSCQKIQLQRINNPHMQEVLGLLQYGTVESCBQVNTSETA	120	DR PROSITE: PS50030; RRM; RNP_1; FALSE_NEG				

Qy	61	TELKGPIEVESVPKRIRKLQLTRNTEPHLGOMEVLDLIVYGVVESGEQVNTDSTSA	120
Db	61	VELKQKLETHESVPKQRSKTQIERNTPQLEWVLGQTLAQGTVNTDSTSA	120
Qy	121	VVNVTYSSKDOAQLDNLNGOLENFETKVAYVTPEDMAQNPLOQPGRGR - GLGORGs	179
Db	121	VVNVTYNNERQTROIAKLNGLHOLENLHVKSUTPDEQSv - - GPENGRRGGFGARGA	176
Qy	180	SQGSP - - GSVSQKQPCDPLRPLRVPTQFVGALIGREGATNTKQSKIDVHRKEN	236
Db	177	PRQGSPVTAGAPKQVDPDPLRPLVPTQYVGALIGREGATNTKQSKIDVHRKEN	236
Qy	237	AGAAEKSTTISPEGSAAACKSLLIMKEADODIKTPEPLKLAHHNFGVLGLIKEG	296
Db	237	AGAAEKAIHSPEGSAAACKSLLIMKEADKTDAEPLKLAHHNFGVLGLIKEG	296
Qy	297	RNLKKEQDTDKTISPLQETLTINPERITVKGNETCAKEEPLMKKIRESYENDIA	356
Db	297	RNLKVEQEOTETKTTISPLQDTLYNPERITVKGSSENCKKAQEIMKVRAYENDVA	356
Qy	357	SNNLQAHLPGLNINAOLGPFTSGMPPTSGPPSAMT - - PPYQPEQSETEVHFQFPA	414
Db	357	AMSLQSHLPGLNLAAYGLFPASSNAVPP - - PSSVSGAAPPSEMPPQEQTFVHFQFPA	413
Qy	415	L5VGALIGKQGOHTKQLRSFAGASIKIAPAEPPDAYKVRMVLITGPPEAQFKAOGRIGKJ	474
Db	414	QAVGATIGKQGOHTKQLRSFASAKTIAPEPPDAYKVRMVLITGPPEAQFKAOGRIGKJ	473
Qy	475	KEEFNVSKEPKLEAHRVPFAAGRVTGKGKTVNELQNLUSAAEVVPRDOPPDENDQ	534
Db	474	KEENFTGPKBKEPKLEAHRVPASAAGRVLGKGKTVNELQNLUTAAEVVPRDOPPDNEQ	533
Qy	535	VVKTGHPAICOVAKTQIOLTOVHQHQHQKALQSPSPQRK	579
Db	534	VVKTGHPAISOMAORRTRDIAVQKO-OHOOKS-OSCOLQARK	576

RESULT 7

088477		PRELIMINARY;	PRT;	577 AA.
AC	088477;			
DT	01-NOV-1998	(TREMBLrel. 08 Created)		
DT	01-NOV-1998	(TREMBLrel. 08 Last sequence update)		
DT	01-JUN-2002	(TREMBLrel. 21 Last annotation update)		
DE	Coding region determinant binding protein (Coding region determinant-binding protein).			
DE	IGFBP1 OR CRDPB.			
GN	Mus musculus (Mouse).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butherida; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OC				
OX	NCBI_TaxID=10030;			
RN	[1]			
RP	SEQUENCE FROM N.A.			

RX MEDLINE=92127743; PubMed=1559612;  
RX Bernstein P.L., Herrick D.J., Prokipak R.D., Ross J.;  
RA "Control of c-myc mRNA half-life in vitro by a protein capable of  
RT binding to a coding region stability determinant.";  
RA Genes Dev. 6:642-654 (1992).  
[2]  
RN RP SEQUENCE FROM N.A.  
RN MEDLINE=94158886; PubMed=8114742;  
RX RA Herrick D.J., Ross J.;  
RT RT "The half-life of c-myc mRNA in growing and serum-stimulated cells:  
RT influence of the coding and 3' untranslated regions and role of  
RT ribosome translocation.";  
RT Mol. Cell. Biol. 14:2119-2128 (1994).  
[3]  
RN RP SEQUENCE FROM N.A.  
RX MEDLINE=94179348; PubMed=8132663;  
RX Prokipak R.D., Herrick D.J., Ross J.;  
RA "Purification and properties of a protein that binds to the C-terminus  
RT of human c-myc mRNA.";  
RT "Coding region of human c-myc mRNA.";  
RT J. Biol. Chem. 269:9261-9269 (1994).  
RL

Db	237	AGAAEKAIHVSHPSPECCSACKMILETMHREAKTKTAADEVPLKILAHNNFVGRLIGNEG	296	QY	180	SRGSP--GSVSKQKPCDPLRLLVPTQFGATIGKEGATIRNITKOTOSKIDVHRKEN	236
QY	297	RNLKTEQDDTKITISPLQETLTNPERTITYGVNETCAAKEETMKTRTRESYENDA	356	Db	177	PRQGSAVAGAPAKQQDIPPLRLLVPTQFGATIGKEGATIRNITKOTOSKIDVHRKEN	236
Db	297	RNLKVEQDTEKITISLQDLTNPERTITYGVNETCAAEQIMMKVREAYENDA	356	QY	237	AGAAEKSTTSLSPGTSAAKSILEMIMEKAQDIKEETPLKLIAHNHFVGRLIGEG	296
QY	357	SMNLQHLPGLNLNAQGLPPTSGMPPPSGPSAMT--PPYPQPEQS-ETEVYHQFTP	413	Db	237	AGAAEKASVHSPECCSACMILEMMKEARDKTDAEVLKTLAHNFVGRLIGEG	296
Db	357	AMSLOHLPLGLNLAAVOLPFASSAPP--PPSSVGAAYSSMOPAQENQVQTP	413	QY	297	RNLKIEQDIDTKITISPLQELTLYNPERTITYGVNETCAAEQIMMKIRESYENDA	356
QY	414	ALSGVATIGKGOKHOIKOLSREGASTKIAPEAPDAKVRMVITGPPEAQPKAQGRIVG	473	Db	297	RNLKVEQDTEKTISLQDITLYNPERTITYGVNETCAAEQIMMKVREAYENDA	356
Db	414	AQAVGALLIGKGOKHOIKOIKSREGASTKIAPEAPDAKVRMVITGPPEAQPKAQGRIVG	473	QY	357	SMNLQHLPGLNLNAQGLPPTSGMPPPSGPSAMT--PPYPQPEQS-ETEVYHQFTP	413
QY	474	IKEENFVSPEEKVLEAHNTIRVPSFARGVIGKGKTVNELQNLSSAEVVYPRDOPDEND	533	Db	357	AMSLQHLPGLNLAAVOLPFASSAPP--PPSSVGAAYSSMOPAQENQVQTP	413
Db	474	IKEENFGPREEVLEHTIRVPSAQRVIGKGKTVNELQNLTAEEVVYPRDOPDEND	533	QY	414	ALSGVATIGKGOKHOIKOIKSREGASTKIAPEADAKVRMVITGPPEAQPKAQGRIVG	473
Qy	534	QVVKITGHYACQVAKQKIQEILTOKHOHQOQKALQSGPPSRRK	579	Db	414	AQAVGAIIGKGOKHOIKOIKSREGASTKIAPEADTDSKVMTVITGPPEAQPKAQGRIVG	473
Db	534	QVVKITGHYACQVAKQKIQEILTOKHOHQOQKALQSGPPSRRK	579	QY	474	IKEENFVSPEEKVLEAHNTIRVPSFARGVIGKGKTVNELQNLSSAEVVYPRDOPDEND	533
Db	534	QVVKITGHYACQVAKQKIQEILTOKHOHQOQKALQSGPPSRRK	579	Db	474	LKEENFGPREEVLEHTIRVPSAQRVIGKGKTVNELQNLTAEEVVYPRDOPDEND	533
<b>RESULT 8</b>							
Q9NZI8		PRELIMINARY;	PRT;	577 AA.	Q9Y6M1	PRELIMINARY;	PRT;
AC	Q9NZI8				Q9Y6M1		
AC	Q9NZI8				AC		
DT	01-OCT-2000	(TREMBLrel. 15, Created)			DT	01-NOV-1999	(TREMBLrel. 12, Created)
DT	01-OCT-2000	(TREMBLrel. 15, Last sequence update)			DT	01-NOV-1999	(TREMBLrel. 12, Last sequence update)
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)			DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)
DE		DE			DE		
OS	Homo sapiens (Human)	DE-pinding protein CRDBP.			GN		
RA	Ioannidis P., Trangas T., Dimitriadis E., Samiotaki M.,				OS	Homo sapiens (Human)	
RA	Panoutsakopoulou G., Kyriazoglou I., Voutzoukis S., Tsiaialis C.M.,				OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
RA	Kittas C., Agnantis N., Pandis N.;				OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
RA	"ectopic expression of a K-h domain containing protein, highly homologous to both human IMP-1 and mouse CRD-BP, in benign and malignant mesothelial tumors."				OX		
RL	Submitted OCT-1999 to the EMBL/GenBank/DBJ databases.				RN		
DR	ENBIL: AF1098254; AAC37203.1;				RP	SEQUENCE FROM N.A.	
DR	DR MEDLINE=9207072; PubMed=10190901;				RX		
DR	DR InterPro: IPI004087; KH_dom.				RA	RA	
DR	DR InterPro: IPI004088; KH_type_1.				RT	"A novel cytoplasmic protein with RNA-binding motifs is an autoantigen	
DR	DR InterPro: IPI004089; RNA_rec_mot.				RT	in human hepatocellular carcinoma."	
DR	DR PEAK: P00013; KH_domain_4.				RL	J. EXP. MED. 189:1101-1110(1999).	
DR	DR PEAK: PF00076; rrm_2.				DR	EMBL: AF057352; ADD315616.1; -.	
DR	DR SMART: SM00322; KH_4.				DR	InterPro: IPI004087; KH_dom.	
DR	DR SMART: SM00360; RRM_2.				DR	InterPro: IPI004088; KH_type_1.	
DR	DR PROSITE: PS50084; KH_TYPE_1; 4.				DR	InterPro: IPI004089; RNA_rec_mot.	
DR	DR PROSITE: PS50102; RRM_2.				DR	InterPro: IPI004090; KH_type_1.	
SQ	SEQUENCE 577 AA: 63456 MW:	0749A060F252D81D CRC64;			DR	SMART: SM00322; KH_2.	
Query Match	73.8%	Score 2181;	DB 4;	Length 577;	DR	PROSITE: PS50084; KH_TYPE_1; 4.	
Best Local Similarity	73.9%	Score 3.3e-136;	DB 4;	Length 577;	DR	Prosite: PS50102; RRM_2;	1EDFEB100443DDC4 CRC64;
Matches	433;	Conservative	61;	Mismatches 76;	Indels 16;	Gaps 8;	
QY	1	ANKLYGNLSENAAPSDELSIFDKAIPKVSGPPELVKIGYAFYVDCPDESALKAFALSGK	60	QY	1	MNKLYGNLSENAAPSDELSIFDKAIPKVSGPPELVKIGYAFYVDCPDESALKAFALSGK	60
QY	1	ANKLYGNLSENAVPSDELSIFDKAIPKVSGPPELVKIGYAFYVDCPDESALKAFALSGK	60	Db	2	MNKLYGNLSENAAPSDELSIFDKAIPKVSGPPELVKIGYAFYVDCPDESALKAFALSGK	61
Db	1	VELQGRLEIHSVPKORIRQLTENIPPHOMEVLSLQYQGYVESEQWNTDSEA	120	QY	61	TELHKRPTEVEHSVPKORIRQLTENIPPHOMEVLSLQYQGYVESEQWNTDSEA	120
Db	61	VELQGRLEIHSVPKORIRQLTENIPPHOMEVLSLQYQGYVESEQWNTDSEA	120	Db	62	VELHGRKMEVDYSRKRSRQIQRNPHQLEVLDGGLLAQYGVNVEQVNIDETA	121
QY	121	VVNVTTSSKQARQDARLNGRQLENFTKVKAYTPDEMAAQONPLQQRGRR-GIGQRGS	179	QY			
Db	121	VVNVTTSSKQARQDARLNGRQLENFTKVKAYTPDEMAAQONPLQQRGRR-GIGQRGS	179	Db			

QY	121	VNVNTYSKSDQARQDQKINGFOLLENFLKVAYIPDEMAQONPQQPERRGIGORSS	RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,	
DB	122	VNVNTYATREPAKTAEMEKSQHOPENSYKISKIPDEEVSSSPPO -RQR -GDHSSR	RA	Li X., Mattei C., McLeod T.C., McPherson D.,	
QY	181	RQG -SPGSVSKOPCDPLRLYPTOFGAIGKEGATIRNITKOTQSLDVKHKENAGA	RA	Merkulov G., Millsina N.V., Moبارry C., Morris J., Moshrefi A.,	
DB	178	EQGHAGFTGTSQARQDPFRILYPTQFGAIGKEGATIRNITKOTQSLDVKHKENSA	RA	Mount S.M., Moy M., Murphy B., Murphy D.M., Nelson D.L.,	
QY	240	AEKSTLSTSPEGTGSACKSILEMHKAQDIFTKETPLKTLAHNFGRGLICKEGRNL	RA	Nelson D.R., Nixon K.A., Nixon K.K., Pollard J., Purvis V., Reese M.G.,	
DB	238	AEPVTHAPEGTSBACRMILEMQADEKLAEPKTLAHNGRGLICKEGRNL	RA	Reinert K., Remington K., Saunders S., Pollard J., Purvis V., Scheeler F., Shen H.,	
QY	300	KKIEODTDFKTISPLQETLYNPERTIVKGIVVERCAKAFFETKKIRFSYNDIASM	RA	Shue B.C., Sieden-Kiamos I., Simpson M., Skupski M.P., Smith T.,	
DB	298	KVIEHETGKTRISSLQDSIYNPERTIVKGIVVEASATEIMKLRAFENDMLAVN	RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,	
QY	360	LOAHLIPGLNUNALGLFPTSGMPPTSGPPSAMTPPYPOFEQ -SETETVHQFIPALSVG	RA	Swirkas R., Teeter C., Turner R., Venturini E., Wang A.H., Wang X.,	
DB	358	TMS-----GYF -----S-LYPHHOFGPFPHHSYPPQETVNLFIPQAVG	RA	Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbergach J.,	
QY	419	ATIGKOGHTKQLSRAGASIKTAPAAEPADAKYRMVITTGPEAQFKAQGRIGYKIKEEN	RA	Williams S.M., Woodgate T., Worley K.C., Wu D., Yang S., Yao Q.A.,	
DB	398	ATIGKGAHKTKQLARFAGASIKTAPAAEPDVSMRVITTGPEAQFKAQRIEGKLEEN	RA	Yeh J., Yeh R.-F., Zaveri J.S., Zhao Q., Zheng L., Zheng X.H., Zhou F.N., Zhou X., Zhu S., Smith H.O.,	
QY	479	FVSPKREVKELEAHTRVPSAAGRVIGKGGTYNELQUNISSAEVYPRDCTPDENDQVYVK	RA	Gibbs R.A., Myers E.W., Rubin G.M., Venet J.C.,	
DB	458	FPNPKEVKLEAHTRVPSAAGRVIGKGGTYNELQUNLTSAEVIVPRDCTPDENEIVR	RA	RTL Science 287:2185-2195(2000).	
QY	539	ITGHFYACOVAQRKIQELTQVKOHQOC	RA	The genome sequence of <i>Drosophila melanogaster.</i>	
DB	518	ITGHFFASQTAQRKIREIVQVKQOEQQ	RA	RL	
QY	539	ITGHFYACOVAQRKIQELTQVKOHQOC	RA	DR	
DB	518	ITGHFFASQTAQRKIREIVQVKQOEQQ	RA	EMBL: AE003484; AAF47058.1;	
DB	358	TMS-----GYF -----S-LYPHHOFGPFPHHSYPPQETVNLFIPQAVG	RA	DR	
DB	358	TMS-----GYF -----S-LYPHHOFGPFPHHSYPPQETVNLFIPQAVG	RA	InterPro: IPRO04087; KH_dom.	
DB	358	TMS-----GYF -----S-LYPHHOFGPFPHHSYPPQETVNLFIPQAVG	RA	FBgn0030335; Imp.	
DB	358	TMS-----GYF -----S-LYPHHOFGPFPHHSYPPQETVNLFIPQAVG	RA	DR	
DB	358	TMS-----GYF -----S-LYPHHOFGPFPHHSYPPQETVNLFIPQAVG	RA	InterPro: IPRO04088; KH_type_1.	
DB	358	TMS-----GYF -----S-LYPHHOFGPFPHHSYPPQETVNLFIPQAVG	RA	Pfam: PF00013; KH-domain; 4.	
DB	358	TMS-----GYF -----S-LYPHHOFGPFPHHSYPPQETVNLFIPQAVG	RA	SMART: SM00332; KH_4.	
DB	358	TMS-----GYF -----S-LYPHHOFGPFPHHSYPPQETVNLFIPQAVG	RA	PROSITE: PS00084; KH_TYPE_1;	
DB	358	TMS-----GYF -----S-LYPHHOFGPFPHHSYPPQETVNLFIPQAVG	RA	SEQUENCE: 558 AA; 6748 MW; 833C9C56954C139C CRC64;	
QY	121	VNVNTYSSKDQARQALDK-----LNGFQLENFTLKAYIPDEMAAQNPQLQOPRGRGLGQ	RA	Query Match: 28/38; Score: 835.5; DB: 5; Length: 558;	
DB	14	ISNNYYQQQSLRILRYDAAVNGEYFEG -----SKLRAQDLDENQRRESQR	RA	Best Local Similarity: 38.5%; Pred. No. 3.6e-47; Indels: 147; Gaps: 12;	
DB	14	ISNNYYQQQSLRILRYDAAVNGEYFEG -----SKLRAQDLDENQRRESQR	RA	Matches: 198; Conservative: 86; Mismatches: 147;	
QY	121	VNVNTYSSKDQARQALDK-----LNGFQLENFTLKAYIPDEMAAQNPQLQOPRGRGLGQ	RA	QY 121 VNVNTYSSKDQARQALDK-----LNGFQLENFTLKAYIPDEMAAQNPQLQOPRGRGLGQ 176	
DB	14	ISNNYYQQQSLRILRYDAAVNGEYFEG -----SKLRAQDLDENQRRESQR	RA	Best Local Similarity: 38.5%; Pred. No. 3.6e-47; Indels: 147; Gaps: 12;	
DB	14	ISNNYYQQQSLRILRYDAAVNGEYFEG -----SKLRAQDLDENQRRESQR	RA	Matches: 198; Conservative: 86; Mismatches: 147;	
QY	177	RGSSROGSPGSVSKQKPCDPLRLYPTQFGAIGKEGATIRNITKOTQSLDVKHKEN	RA	QY 177 RGSSROGSPGSVSKQKPCDPLRLYPTQFGAIGKEGATIRNITKOTQSLDVKHKEN 236	
DB	61	-NQNPYPGMPGPGQDADPRLYQSEPMVQGAIQRGQStiritoQSRdYDHREN	RA	DB 61 -NQNPYPGMPGPGQDADPRLYQSEPMVQGAIQRGQStiritoQSRdYDHREN 118	
QY	237	AGAAFKSTILSPEGTGSAACKSILEMHKAQDQIKTE -----EPIKLIKLAHHNNEGR	RA	QY 237 AGAAFKSTILSPEGTGSAACKSILEMHKAQDQIKTE -----EPIKLIKLAHHNNEGR 290	
DB	119	VGSLERKSIITYGNENCYNACRKLLEMQAEISTNGKELSPESCEICLKILAHNNLGR	RA	DB 119 VGSLERKSIITYGNENCYNACRKLLEMQAEISTNGKELSPESCEICLKILAHNNLGR 178	
QY	291	LIGKGRNLKKIEQDDTDTKITSPLQELTLYNPERTIVKGIVNETCAKAEELMKIKRES	RA	QY 291 LIGKGRNLKKIEQDDTDTKITSPLQELTLYNPERTIVKGIVNETCAKAEELMKIKRES 350	
DB	179	IIGKSENTRKIMODTDTKITYSSINDINSPLNLERITYVGLENIENSRAENQISTKURQ	RA	DB 179 IIGKSENTRKIMODTDTKITYSSINDINSPLNLERITYVGLENIENSRAENQISTKURQ 238	
QY	351	YENDIASMNLQAHLIPQNLNALGLFPTSCMPPPTSGP -----PEAMTUPP	RA	QY 351 YENDIASMNLQAHLIPQNLNALGLFPTSCMPPPTSGP -----PEAMTUPP 398	
DB	239	YENDQAMAFQSLMPGPQGAVFNPMGPQGHPMAN-MSPGPNGAVFNPMGPQGHPMAN	RA	DB 239 YENDQAMAFQSLMPGPQGAVFNPMGPQGHPMAN-MSPGPNGAVFNPMGPQGHPMAN 296	
QY	399	QEQQSETEVYHOFIPALSVGAIIGKQGQHQLRSFAGASTIKAPAEA ---EDAKYRMV	RA	QY 399 QEQQSETEVYHOFIPALSVGAIIGKQGQHQLRSFAGASTIKAPAEA ---EDAKYRMV 454	
DB	297	-FPNDIQTETYLYTPNNAVGATIGTRSHIRSIMRSNAILDPAQPLDQOTEK	RA	DB 297 -FPNDIQTETYLYTPNNAVGATIGTRSHIRSIMRSNAILDPAQPLDQOTEK 355	
RA	13	(TREMBLrel. 13, Created)	RA	RA Amanatidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.P., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Worms J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazquez R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abri J.F., Agbayani A., An H.-J., Andrews P.J., Baldwin C., Bailew R.M., Basu A., Baxendale J., Bayraktagoglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhansali D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Bröttner P., Burris K.C., Busam D.A., Butler H., Cadieu E., Cencera A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doucet L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evans C.J., Ferreira C., Ferreira S., Fleischmann W., Foster C., Gabrilian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Haris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lai K.,	RESULT 11
RA	15	(TREMBLrel. 15, Last annotation update)	RA	RA ID Q9NGS9, PRELIMINARY; PRT; 566 AA.	
RA	15	(TREMBLrel. 15, Last sequence update)	RA	RA AC Q2NGS9, AC Q2NGS9, PRELIMINARY; PRT; 566 AA.	
RA	15	(TREMBLrel. 15, Last sequence update)	RA	RA DT 01-OCT-2000 (TREMBLrel. 15, Created)	
RA	15	(TREMBLrel. 15, Last sequence update)	RA	RA DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)	

DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)		
DE	Igf-II mRNA-binding protein	(SP07045P).		
GN	IMP OR CG1691.			
OS	Drosophila melanogaster (Fruit fly)			
EC	Drosophila; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydriidae; Drosophilidae; Drosophila.			
OX	NCBI_TaxID:7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Nielsen J., Nielsen F.C., Christiansen J.: Cloning and Expression of a Drosophila Homolog of IMP/Yg1-RBP.";			
RT	"Cloning and Expression of a Drosophila Homolog of IMP/Yg1-RBP."			
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJU databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J., Champine M., Chavez C., Dorsett V., Farfan D., Frise E., George R., Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J., Nanno J., Paragis V., Park S., Phouanavong S., Wan K., Yau C., Lewis S.E., Rubin B.M., Celniker S.; Submitted (DEC-2001) to the EMBL/GenBank/DBJU databases.			
DR	EMBL; AF241237; AAC63311; -;			
DR	AY059821; AAI39061; -;			
DR	FLYBase; FBgn0030235; Imp			
DR	InterPro; IPRO00487; KH_dom.			
DR	InterPro; IPRO04088; KH_type_1.			
DR	PIR; PF00013; KH_domain; 4.			
DR	PROSITE; PS50084; KH_4_TYPE_1; 4.			
SQ	SEQUENCE 566 AA; 62129 MW; B4C74C5C9B0C830 CRC64;			
Query Match	28 28; Score 835; DB 5; Length 566;			
Best Local Similarity	38.8%; Pred. No. 3.9e-47;			
Matches	197; Conservative 86; Mismatches 147; Indels 78; Gaps 12;			
Qy	121 VVNTYSSKQDARQLDK---LNGFOLENTFLTKVAYIPDEMAAQONPLQPRGRGLSQ 176	:		
Db	14 ISNNYYQQSLLRILDAVYLNGFEGF-----SKHAEQDKNRRRSQ--- 60	:		
Qy	177 RGSSRQGSGSVSKQPKDPLGLRLLVPTQFGVAILKGKEATINIKTQSKIDVHRKEN 236	:		
Db	61 --NQRNPYGMGPGRQADFPLRILVQSEAVGATIGRQOSTTRITQSRARVDVHRKEN 118	:		
Qy	237 AGAAEKSTLSTPEGTSAACKSLIMMKHEAQDIFTEPIKPLKILAHNNVGRJLKEG 296	:		
Db	119 VGSLEKSTYGNFENCTNACERLVMMQAISIN-KGEICLKILAHNNLIGTRIGSK 177	:		
Qy	297 RNLKKEQDPTDKTISPLQELTYNPERTITVKGNVETCAKEEEIMKKIRESYENDIA 356	:		
Db	178 NTIKRIMQDPTDKTIVSINDINSFEDERTITVKGLIEMSRAENQTSKLRQSENDLQ 237	:		
Qy	357 SMNLOAHLPLGLNUNALGLFPETSGMPPTSGP-----PSAMTPPYQFEQSE 404	:		
Db	238 AMAP0SLMPGLIUPMAM-MSTPGNGAVFNTSMPPCSOFSAMSKTPASSVVPV--FPNDL 294	:		
Qy	405 TETVHQFIPALSGAIGKOGOHIKOLSRAGASIKAPAEA---PDAKVRMVLITGGP 460	:		
Db	295 QETTYLYIPNNAVGAIITGRSHIRSIMRFSNASLKIAPLDADKPLDQTERKVITGTP 354	:		
Qy	461 EAQFQAQGRIGYKTKKENEVSPKEEYKLEAHIRPSFAAGRIVGKGKTWNEQNLSSAE 520	:		
Db	355 EGQWKAQYMFMRRECFMCCDTDYRLTVELLASSQVGRIGKGQVNRELQRTGGSV 414	:		
Qy	521 VVPR-----DQTPDENDQVVVKITGHFAYCQVAQRKIQEIL----- 557	:		
Db	415 IKLPHHALAPPSSGDETP-----VHIGLFYSVQSAQRRTIRAMMLSTNPPPTKKQK 467	:		
Qy	558 --TQVKHOQOKA-----LQSGPPRQS 576	:		
Db	468 AAKEQOLQQQSLAGAASSGQQQFQS 495	:		
RESULT	12			
Q9D054	PRELIMINARY; PRIT; 100 AA.	ID	Q9D054	PRELIMINARY; PRIT; 854 AA.
AC		ID	Q9D054	
DT	01-JUN-2001 (TREMBLrel. 17, Created)	DT	01-JUN-2001 (TREMBLrel. 01, Last sequence update)	
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)	DT	01-JUN-2002 (TREMBLrel. 01, Last sequence update)	
DE	TGF-β111Rik protein	DE	TGF-β111Rik protein	
GN	TGF-β2P3 OR 26.01011Rik.	GN	TGF-β2P3 OR 26.01011Rik.	
OS	Mus musculus (Mouse)	OS	Mus musculus (Mouse)	
OC	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OC	Bukaryota; Metazoa; Chromodorea; Rhabditida; Rhabditidae; Rhabditidae; Rhabditidae; Peioderinae; Caenorhabditis.	
NCBI_TaxID	10090; [1]	NCBI_TaxID	6239; [1]	
RN	SEQUENCE FROM N.A.	RN	SEQUENCE FROM N.A.	
RC	STRAIN=C57BL/6J; TISSUE=EMBRYO;	RC	STRAIN=C57BL/6J; TISSUE=EMBRYO;	
RX	MEDLINE=21085660; PubMed=11211851;	RX	MEDLINE=21085660; PubMed=11211851;	
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Hara A., Fukuda S., Kurozawa T., Iwasa M., Nishi K., Kijosawa H., Kondo S., Adachi J., Yamanaka I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batzler S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuno Y., Nikaido I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Botjung C., Carninci P., de Bonaldo M.F., Brownstein M.J., Built C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hall D., Hofmann M., Hume D.A., Kamiy M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P., Nordone P., Ring B., Rinewald M., Rodriguez I., Saramoto N., Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F., Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshioka K., Hasegawa Y., Kohetsuki S., Hayashizaki Y.	RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Hara A., Fukuda S., Kurozawa T., Iwasa M., Nishi K., Kijosawa H., Kondo S., Adachi J., Yamanaka I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batzler S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuno Y., Nikaido I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Botjung C., Carninci P., de Bonaldo M.F., Brownstein M.J., Built C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hall D., Hofmann M., Hume D.A., Kamiy M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P., Nordone P., Ring B., Rinewald M., Rodriguez I., Saramoto N., Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F., Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshioka K., Hasegawa Y., Kohetsuki S., Hayashizaki Y.	
RA	Gustincich S., Hall D., Hofmann M., Hume D.A., Kamiy M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P., Nordone P., Ring B., Rinewald M., Rodriguez I., Saramoto N., Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F., Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshioka K., Hasegawa Y., Kohetsuki S., Hayashizaki Y.	RA	Gustincich S., Hall D., Hofmann M., Hume D.A., Kamiy M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P., Nordone P., Ring B., Rinewald M., Rodriguez I., Saramoto N., Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F., Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshioka K., Hasegawa Y., Kohetsuki S., Hayashizaki Y.	
RA	Nature 409:685-690 (2001).	RA	Nature 409:685-690 (2001).	
DR	AK011797; BAB27848.; -	DR	AK011797; BAB27848.; -	
DR	MGD; MG1:1803539; I912B2P3.	DR	MGD; MG1:1803539; I912B2P3.	
DR	InterPro; IPR000040; RNA_rec_mot.	DR	InterPro; IPR000040; RNA_rec_mot.	
DR	PFam; PF00076; rrm_1.	DR	PFam; PF00076; rrm_1.	
DR	SMART; SM00360; RRM_1.	DR	SMART; SM00360; RRM_1.	
DR	PROSITE; PS50102; RRM_1.	DR	PROSITE; PS50102; RRM_1.	
SQ	SEQUENCE 100 AA; 11249 MW;	SQ	SEQUENCE 100 AA; 11249 MW;	
RT	"Functional annotation of a full-length mouse cDNA collection.";	RT	"Functional annotation of a full-length mouse cDNA collection.";	
RL	RL	RL	RL	
DR	Best Local Similarity 88.7%; Pred. No. 3.1e-24; Indels 0; Gaps 0;	DR	Best Local Similarity 88.7%; Pred. No. 3.1e-24; Indels 0; Gaps 0;	
Matches	86; Conservative 86; Mismatches 147; Indels 78; Gaps 12;	Matches	86; Conservative 86; Mismatches 9; Indels 0; Gaps 0;	
Qy	1 MNKLYIGNLSNAAPSLESIFDAKIVPYSGPPLVKIGYAFVDCPDESALKATEALSGK 60	Qy	1 MNKLYIGNLSNAAPSLESIFDAKIVPYSGPPLVKIGYAFVDCPDESALKATEALSGK 60	
Db	1 MNKLYIGNLSNAAPSLESIFDAKIVPYSGPPLVKIGYAFVDCPDESALKATEALSGK 60	Db	1 MNKLYIGNLSNAAPSLESIFDAKIVPYSGPPLVKIGYAFVDCPDESALKATEALSGK 60	
Qy	61 TEIHKPTEEVESVPRQRTRKLQTRNPPHLOWEV 97	Qy	61 TEIHKPTEEVESVPRQRTRKLQTRNPPHLOWEV 97	
Db	61 MELHGPMEVEHSVPRQRTRKLQTRNPPHQWM 97	Db	61 MELHGPMEVEHSVPRQRTRKLQTRNPPHQWM 97	
RESULT	13	RESULT	13	
Q21605		Q21605		
AC		AC		
DT	01-NOV-1996 (TREMBLrel. 01, Created)	DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)	
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)	DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)	
DE	M88.5 protein.	DE	M88.5 protein.	
GN		GN		
OS	Caenorhabditis elegans	OS	Caenorhabditis elegans	
OC	Bukaryota; Metazoa; Chromodorea; Rhabditida; Rhabditidae; Rhabditidae; Rhabditidae; Peioderinae; Caenorhabditis.	OC	Bukaryota; Metazoa; Chromodorea; Rhabditida; Rhabditidae; Rhabditidae; Rhabditidae; Peioderinae; Caenorhabditis.	
NCBI_TaxID	5239; [1]	NCBI_TaxID	5239; [1]	
RN	SEQUENCE FROM N.A.	RN	SEQUENCE FROM N.A.	
RA	Sulston J.E.;	RA	Sulston J.E.;	

RL	Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.	DT	01-OCT-2000 (TREMBLE1. 15, Created)
RN	[2]	DT	01-OCT-2000 (TREMBLE1. 15, Last sequence update)
RP	SEQUENCE FROM N.A.	DT	01-JUN-2003 (TREMBLE1. 21, Last annotation update)
MEPDINE=9909613; PubMed=3851916;		DE	
RA	none;	ESTS	D23339 (R0339).
RT	"Genome sequence of the nematode <i>C. elegans</i> : A platform for investigating biology";	OS	Oryza sativa (Rice).
RT	Investigating biology";	OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
RT	Science 282:2012-2018 (1998);	OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
RL	EMBL; Z34802; CAA84338; 1. -.	OC	Ehrhartoideae; Oryzeae; Oryzae.
DR	NCBI_TaxID=4530;	OX	
DR	IPRO04087; KH-domain.	RN	[1]
DR	IPRO04088; KH-type_1.	RP	
PFAM	PF00013; KH-domain; 4.	RC	SEQUENCE FROM N.A.
DR	IPRO04084; KH-type_4.	STRAIN=CV; NIPPONBARE;	
DR	P500084; KH_TYPE_1; 4.	RA	Sasaki T., Matsumoto T., Yamamoto K.;
SEQUENCE	854 AA; 94553 MW; D3B40412353ECC21 CRC64;	RT	Os <i>Oryza sativa</i> nipponbare(GA3) genomic DNA, chromosome 1, PAC
Query Match	14.28%; Score 418.5; DB 5; Length 854;	RT	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
Best Local Similarity	24.18%; Pred. No. 2.6e-19;	RL	
Matches	160; Conservative 103; Mismatches 215; Indels 187; Gaps 27;	DR	AP001535; BA92910.1; -.
Qy	78 QRIKQLQIRINPPHQLQEVIDLSLIVQGVVESCEOVNTEAVNN-----VTYSSK 129	DR	InterPro: IPR004087; KH dom.
Db	167 QMOMQA-QQQAQHMHQMQHQQMHQMLPQHQQMHQMQHAQGQYHPHQONQHQA-GHOOSHHSQO 221	DR	InterPro: IPR000088; KH-type_1.
Qy	130 DQARQLDKINGFOLLENFLTRVAYITDPEMA-AQO--NPLQQPRGRGRLQQRGS-SRQGSP 185	DR	SMART: PF00013; KH-domain; 4.
Db	222 QMOMQA-QQQAQHMHQMQHQQMHQMLPQHQQMHQMQHAQGQYHPHQONQHQA-GHOOSHHSQO 279	DR	PROSITE; PS500084; KH_TYPE_1; 4.
Qy	186 G-----SYSKQKP-----C-----DLPLRLVPTOFVGAIIGKEGATIRNTIKQ 224	DR	PROSITE; PS500084; KH_TYPE_1; 4.
Db	280 NHQHNRHNQSHSGPHHIPONLNMMRCMLRDPWTRCVVERSKYHAWIIGPNGSTIKDIASS 339	DR	SEQUENCE 774 AA; 84066 MW; 5992CGFE0F74BB56 CRC64;
Qy	225 TOSKID--VHRKENA--GAAEKSTTISLPEGTSAAACKSILEMHKEA-QDIKFTEEL 277	Query Match	8.9%; Score 262.5; DB 10;
Db	340 TRCRDEVNLSSKERTVGLNDRILTVHGVAEQATKAVARILDYIOSEAKDVNVNGADT 399	Best Local Similarity	21.8%; Pred. No. 4.8e-09;
Qy	278 PLKLAHHNNFTVGRGLIKEGENRLKKTEQDTEKTITISLPOB-----LT-----LYNPERT 326	Matches	129; Mismatches 18; Indels 205; Gaps 20;
Db	400 VLRMRAHNLQCLGRGLIKAGESSKIEIMOKTGNTNTVTKYIEPGGISGLTANELLGLMERT 459	Qy	156 DEMAAQQNPLQQPRGRGRLQGRGSSRQGSPGSVSKOPCDLP-----LRLVLP 203
Qy	327 ITVKG-NVEVKAKABEEIMMKIRESYENDIASMNIQAHLLPGLNLNALGLFPPTSGMPP- 384	Db	57 EDMASRSPSKRP-----FORNSSEONGRKWQTKHNSLQQPLQVQGPVPIRLCP 110
Db	460 IMVRGPSTIEAVVQARALISKLKCYESD_SQLAQSMCOP-----MPPM-MPPI 508	Qy	204 TQFGATIGKEGATIRNTIKQTOSKI-----DVHRKENA 237
Qy	385 -PTSGPPSAMTPPY-----PQFQESTETVYHO-----	Db	111 TSKSGNVIKGGGIIIAKTRQEGVKFRDEVPGCDERIVITAIKDREVSHVOTKEND 170
Db	509 LPPGASSSA(SAPHI)PTPVGMOVQHETASQHVHQNANSFLQPGVLQIQPGTTNLRQ 568	Qy	238 GA-----AEKSI-----244
Qy	411 --FIPALSYGAIIGKQGHQKLSRFAGASAKTAPAEPAVDAY-----	Db	171 GGACSYDVGHNHGEKDTIKEKEDESNEEMDSEKDKDLEKDHDKEDRDESGKDN 230
Db	569 VRMWYVDSMAGALICAKGRNMKMTIRDTCASVKL---BAPEENTQREAEAEEKRKLDET 625	Qy	245 -----TILTSPGTSAACKSITLEIMHKBAQDQIKFTEE-----276
Qy	452 -----	Db	231 DKKADDSSVAKDTNSEPAOLEKMPPLAYKAILVF----DRIFVNMEMGTDGASG 285
Db	626 DSGCEGVASGDHPQEFLEDNATINSSDALEKPKPVSERVMTINGDDLQLQKAQSYYFSK 685	Qy	277 ---IPKILAHNNFTVGRGLIKEGRNKKLIEQDTDKTTISPLQELTLYNPERTIVKGN 332
Qy	474 IKEENFVSP-----KEEVKLFEDIHRYPSFAAGRVIGKGSTVNELQNLSSAENVYVPR 525	Db	286 ERNHVSRLLVLDQSOVGLLGKNGSVTKQMSTDSCIRVS-----KDK 329
Db	686 IAETSSLPSGMGDRSHMLRITEVSYPTTRIGRIGKGGVNREQRITGAVVXKPE 745	Qy	333 VENCAKABEEIMMKIRE-----SYNDIASMNIQAHL-----IPGLNLNALGLP--PPTS-----380
Qy	526 DQTPD-----ENDQYVVKITGHFYACQVAQRKIQEILTOVKQ--HQOQKA-LQS 571	Db	330 LPICALDRDELCOILHKKITYSLHSFLQTGELDSVRLQNLTVQALLLTHPPKESDVL 389
Db	746 EERNGGEVYRHDDLEEDMTMIRTIGNNYSTMHNQFRALAHLYNEYRSQGDHRNKSSDYKG 805	Qy	381 -----GMPP----PTSGPPSAMTPPYQ-----FEQSETETVH 409
Qy	572 GPPQS 576	Db	390 GAHNGGSSRSFFNQPDYLPPGQPNLHLPPCGPMVNAHLNPPEALMHGHSVPPEPLTFR 449
Db	806 GRPHS 810	Qy	410 QFTPALSYGAIIGKQGHQKLSRFAGASAKTAPAEPAVDAY-----EAQFKKA 466
RESULT 14	Q9L128	PRELIMINARY;	PRT; 774 AA.
ID	Q9L128	PRELIMINARY;	PRT; 542 AA.
AC	Q9L128;	PRELIMINARY;	PRT; 542 AA.

DT 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Putative nucleic acid binding protein.  
 GN OSINBB0048D20\_13.  
 OS Oryza sativa (Rice).  
 OC Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TAXID=4510;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Cv. NIPPONBARE;  
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,  
 RA Gansberger K., Brenner M., Burgess S., Hance M., Shvartebeyn M.,  
 RA Tsirrin F., Riggs F., Hsiao J., Blunt S., Pai G.,  
 RA Vanaken S.E., Utterback T.R., Feldblum T.V., Kalb E., Quackenbush J.,  
 RA Salzberg S.L., White O., Fraser C.M.,  
 RT "Oryza sativa" chromosome 3 BAC OSINBB0048D20 genomic sequence.;"  
 RL Submitted (FB-2002) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AC084766; AANL8259; 1  
 SQ SEQUENCE 542 AA; 57843 MW; B28A5-BB2450F336 CRC64;

Query Match	Score	Length
Best Local Similarity	8.8%	542;
Matches	22.5%	Pred. No. 3.9e-09;
Matches	94;	Mismatches 191;
		Indels 77;
		Gaps 17;

Qy 168 PRGRGLGQR-----GSRSRQSPG-SVSKORPCDDPLRLVPTQEVGLIGKEGATIR 219  
 Db 4 PGHRNSHKGKHSYDAENGGKRRNPGDDTYAQPDDTVYRYLCPSRKIGSLIGRGGEIAK 63

Qy 220 NITKQTOSKRIDYHRKENAGAAERSITILSTPESTAACKSILEIMHK--EAQDIKF--- 273

Db 64 QMRADTQKIRI\_GESVQCDERVITFSSSRETIN---TLYDAEYKVCPADALERVHE 118

Qy 274 -----TEEPIPKILAHNNFVGRLIGKEGRNLKKIEQDTDKITISPLQE 317

Db 119 KLSDDDCNEESDEGLAQVTVRLLVPDSQIGEJGKHIIQGIRSDTGIAHIVLSNEN 178

Qy 318 L-TLYNPERTIVKGNTETCAKEEEIMKKIRESYENDIASMLQA-----HL 364

Db 179 LPACAISGBELLQLSGDSTVVRKALLQVSRL--HONPSRSQHLLASSMTQPYPGSHL 235

Qy 365 IPGMLNLAUGLFPPTS--GMPPTPSGPPSAMTPPYPOFEQSETETVHQFTALPSVGAI 421

Db 236 GSSSTAPVQGITPLISSYGYKGDVAWPSIXQPRRESSAKFESURLCASNYGGV 295

Qy 422 GKQGQHIIKOLSRAGASIKIAPAEAPAKVRMVTITGPPEAOFAKGRTYGRIKEENEV 481

Db 296 GKGGIIGIKQRQESGAFIVKDSSNTEDD--CIIIVSAKEFF-EDPVSPINAAYHLQ 349

Qy 482 PKEVKLEHIRVSEAA-----GVIGKGGKTYNELQNISSAEV-VYPRDQTDP- 530

Db 350 PRCEKTDPESAIPSYTTRLLVSTSRICGLGGSITTEIRTSRANIRLSKENVPKV 409

Qy 531 -ENDOVVVKITGHYACQVQAQRKIQEILTVKOH--OOQKALOSGPP 574

Db 410 PAEEDEEMVQISGDL--DVVRHALIQTTRKLNFEEFRGALSFGPP 453

Search completed: April 24, 2003, 16:14:14  
 Job time : 42 secs

Scoring table:	BLOSUM62		
Gapop:	10.0	Gapext:	0.5
Searched:	112892 seqs,	41476328 residues	
Total number of hits satisfying chosen parameters:	112892		
Minimum DB seq length:	0		
Maximum DB seq length:	2000000000		
Post-processing:	Minimum Match 0%		
	Maximum Match 100%		
	Listing first 45 summaries		
Database :	SwissProt_40.0		
Pred.	No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.		
		RESULTS	
Result No.	Score	Query Match Length	Description
1	227	7.7	510 1 NOA_HUMAN P51513 homo sapien
2	224.5	7.6	339 1 PCB3_HUMAN P57721 homo sapien
3	219.5	7.4	339 1 PCB3_MOUSE P57722 mus musculus
4	213.5	7.2	403 1 PCB4_HUMAN P57723 homo sapien
5	210.5	7.2	403 1 PCB4_MOUSE P57724 mus musculus
6	210.5	7.1	606 1 Y475_ARATH P58223 arabidopsis
7	209.5	7.1	356 1 PCB1_HUMAN Q15365 homo sapien
8	209.5	7.1	356 1 PCB1_RABBIT Q19048 orctyolagrus
9	207.5	7.0	1268 1 VGLN_HUMAN Q03341 homo sapien
10	204	6.9	413 1 PBP2_YEAST P38151 saccharomyces
11	202.5	6.9	365 1 PCB2_HUMAN Q15366 homo sapien
12	202	6.8	362 1 PCB2_MOUSE Q61990 mus musculus
13	194.5	6.6	1270 1 VGLN_CHICK P81021 gallus gallus
14	190	6.4	463 1 ROK_HUMAN Q07244 homo sapien
15	190	6.4	463 1 ROK_RABBIT Q19049 orctyolagrus
16	190	6.4	464 1 ROK_MOUSE Q60577 mus musculus
17	183	6.2	1222 1 S160_YEAST P06105 saccharomyces
18	170.5	5.8	629 1 PAB2_ARATH P42731 arabidopsis
19	164.5	5.6	644 1 PAB1_HUMAN Q13310 homo sapien
20	160.5	5.4	636 1 PAB1_MOUSE Q29341 mus musculus
21	159.5	5.4	470 1 ROK_HUMAN Q15233 homo sapien
22	159.5	5.4	463 1 PAB1_RABBIT Q11940 homo sapien
23	156	5.3	633 1 PABP_XENLA P20965 xenopus laevis
24	152.5	5.2	381 1 YBD2_YEAST P38195 saccharomyces
25	150	5.1	344 1 SERF_HUMAN Q13247 homo sapien
26	147.5	5.0	747 1 FCA_ARATH Q04425 arabidopsis
27	145.5	4.9	414 1 NSRL_YEAST P7476 saccharomyces
28	145.5	4.9	489 1 SFRP_MOUSE Q8ve97 mus musculus
29	145	4.9	403 1 NOP3_YEAST Q15560 saccharomyces
30	139.5	4.7	522 1 PAB2_HUMAN Q15097 homo sapien
31	138.5	4.7	424 1 S3B4_HUMAN Q15427 homo sapien
32	135.5	4.6	359 1 BLV2_HUMAN Q12926 homo sapien
33	135.5	4.6	360 1 ELV2_MOUSE Q60899 mus musculus

DR	MTM_602157; InterPro: I PRO04087; KH_dom.	RA Makeyev A.V.; Liebhaber S.A.; Identification of two novel mammalian genes establishes a subfamily of KH-domain RNA-binding proteins.;
DR	InterPro: I PRO04088; KH_main_type_1.	RT
DR	PFAM: PF00013; KH_domain; 3.	RT
DR	SMART: SM00322; KH; 3.	RL Genomics 67:301-316(2000).
DR	PROSITE: PS50084; KH_TYPE_1; 3.	CC FUNCTION: SINGLE-STRANDED NUCLEIC ACID BINDING PROTEIN THAT BINDS PREFERENTIALLY TO OLIGO DC (BY SIMILARITY).
DR	Nuclear protein; RNA-binding; Repeat; Alternative splicing; Antigen.	CC SUBCELLULAR LOCATION: Nuclear (By similarity).
KW		CC -!- SIMILARITY: CONTAINS 3 KH DOMAINS.
FT DOMAIN	52	CC
FT DOMAIN	212	CC
FT DOMAIN	212	CC
FT VARSPLIC	427	CC
FT VARSPLIC	462	CC
FT VARSPLIC	88	CC MISSING (IN TUMOR ISOFORM).
FT VARSPLIC	90	CC MISSING (IN ISOFORM 3).
FT VARSPLIC	153	CC MISSING (IN TUMOR ISOFORM 3).
FT VARSPLIC	176	CC VTIIVNS -> KININIS (IN TUMOR ISOFORM).
FT VARSPLIC	184	CC
FT VARSPLIC	185	CC MISSING (IN TUMOR ISOFORM).
FT VARSPLIC	510	CC MISSING (IN TUMOR ISOFORM).
SEQUENCE	510 AA;	CC 52055 MW; 7B9BA8B67 CAB6984 CRC44;
Query Match	7.7%;	Score 227; DB 1; Length 510;
Best Local Similarity	22.4%;	Pred. No. 5.1e-07;
Matches	119; Conservative 77; Mismatches 179; Indels 156; Gaps 20;	
Qy	155 PDEMAQQNPLQQPRGRGLGORGSSQGSPGSVSKQPCDIPFLRLLVPTPFGATIGE 214	DR HSSP: Q07244; AAG09240.1; -.
Db	23 ED--SKRPLEAP----PEAGSKTRNTNGEDQYF---LKVLPSYAAGSTIGK 68	DR HGNC: HGNC_8651; PCBPs 3.
Qy	215 GATIRNITKOTQSKIDVHRKENA---GAAEKSTILSTPGTSACKSLEIMHKEAQ 269	DR InterPro: I PR00404087; KH_dom.
Db	69 GOTIVQLQTKTGTAVTKLKLSSKDFYPTERYCLQQTVEALNAVGYFAEKI REMPO 128	DR InterPro: I PR00404088; KH_type_1.
Qy	270 DIKFTEIPL-----KILAHNNFVGR-----KILAHNNFVGR-----KILAHNNFVGR 290	DR Pfam: SM00022; KH; 3.
Db	129 NYAKTPPVSVLQPQTWNPDRIKQTLESSLSPPTTKSSPDPMITSRANQVKLIVPISTAGL 188	DR PROSITE: PS50084; KH_TYPE_1; 3.
Qy	291 LIGKEGRNLKKIEQDTDKITISPLQETLTYNPERITTYGNVETCAKAEEEIMKIRE- 349	DR Nuclear protein; RNA-binding; Ribonucleoprotein; DNA-binding; KW Repeat.
Db	189 IIGKGCATVAVMQSQGAWVQLSQKPP-GINLQREVVTVSGEBQNKAVALIIOIQD 247	DR DOMAIN 13 KH 1.
Qy	350 -----SYEN---DIASMN-----LQAHLIPGUNLNNALGLEPPT- 379	DR DOMAIN 97 150 KH 2.
Db	248 PQGSSCLNISANYTGAVANSNPTGSSYANTAEVLPTAAAAGLHNLAGVAAPYL 307	DR DOMAIN 261 131 KH 3.
Qy	380 SGMPPTSGPPSAMTPPPQFQESETFETVHOIPTALSGAIIKGKOHIKOLSRPAGASI 439	DR SEQUENCE 339 AA; 35938 MW; FEF19FF7590C4188 CRC64;
Db	308 SGF---TGNDLVAT-----SAINTL-----ASGYVNNTLGLLSQAAA-TGALA 349	DR
Qy	440 KIAAPAPDAKVRNAYITGPPAQKFAQRTYGIK-----474	DR
Db	350 AAAASNPAPAAANLARYASEA-----SISGSGTAGGTAGTALGSLAAATAATANGYGAASP 407	DR
Qy	475 -----KEENFISPKKEVKLEAHTRPVSPFAASVGGVIGRKGGKTYNEELNQNSAEEVYVPR- 525	DR
Db	408 LAASALGTEKSTDGSKDV-----IAVPENVLGAIIQGGKTLVEOELGARIQISSK 463.	DR
Qy	526 -DQTPEENDQVVKVITGHYACQVAQKRIQSBILTYKQHOOQQKALQSGPQ 575	DR
Db	464 GEFPVPTTRNKRVT-ITGTPAATQAAQYLQRIT-----YEQGVRAANPQ 507	DR
RESULT 2		DR
PCB3_HUMAN	STANDARD;	DR
AC	PCB3_HUMAN	PRT; 339 AA.
DT	P57721; 16-OCT-2001 (Rel. 40, Created)	DR
DT	16-OCT-2001 (Rel. 40, Last sequence update)	DR
DE	16-OCT-2001 (Rel. 40, Last annotation update)	DR
GN	PCBP3.	AC P57722; 16-OCT-2001 (Rel. 40, Created)
OS	Homo sapiens (Human).	DT 16-OCT-2001 (Rel. 40, Last sequence update)
OC	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Buteleostomi; Muridae; Murinae; Mus;	DE Poly(rC)-binding protein 3 (Alpha-CP3).
OX	NCBI_TaxID:9606;	GN PCBPs
RN	[1]	OS Mus musculus (Mouse).
RP	SEQUENCE FROM N.A.	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
RX	MEDLINE=20396135; PubMed=10936052;	OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;



Db	59 PERITIT- GSTAAVFFAVSMIAFKLDEDCAAPANGNVSRPPVTLLVIPASQCGSL	116	Db	18 LTLRMLMHGKEVGSIIGKKGETVKIREQSSARITI- - -SEGSCPERITIT--GSTAA	71
Qy	292 IGKECRNLLKIEQQTDTITISPLQEITLYNPERTITKGIVETCAKAEEIMMKTREY	351	Qy	256 ACKSTLEIMHEKAODI-----KTEETIPLKILAHNNFGRUETKIEQRT 306	
Db	117 IKGASTKIKIRETGAQVVA -GDLLPNSTERAVTVSGYPDAILLCRQICAVILEP	174	Db	72 VFAVSMIAFKLDEDCAAPANGNVSRPPVTLLVIPASQCGSLIGAKGTKEIRETT	131
Qy	352 ENDIASMNLQAHHLIPGLNLNUNAAGLFPPSGMPPTSGPSAMTP ----- PY	397	Qy	307 DTKITISPLQELTYNPERTITYKGIVETCAAAEELMKTKRESYENDIASMNLQAHHLIP	366
Db	175 PK---GATIPYH- PSLSLGTV- LLSANGF --SVOQGYAGTVAEVTKLQQLSHAYF	226	Db	132 GAQVDA- GDLLENSTERAVTVSGVPDAILLCRQICAVILESPPK -- GATIPYH- P	184
Qy	398 -----POFQ-SETETVQFIALSYGAIIGKQGHIKOLSRFAGASTKIAAPEADPAK	450	Qy	367 GLNLNLALGFPPPTSGMPPTSGPSAMTP----- PY----- POFEQSE	406
Db	227 ATPSVPGDGPQTQSSQFLYDNLIGCVRGQSKSEIRMSGAHKIG-NOAEAGG	285	Db	185 SLSLGNV-LLSANGF- -SVQGQYVPAEVTKLQOLSGHAPVAPSVPGMGPSTOT	241
Qy	451 VRMVLTIGPPEAQFKAQ 467		Qy	407 TYHQF-IPALSVGAIIKGQGHIKQSLRFAGASIKTIAAPEADPAKDAYRMVYTITGPPAEQFK	465
Db	286 ERHVITGSPVSTALQ	302	Db	242 SSQEFVNPNDLIGCVRGQSKSEIRMSGAHKIG-NOAEAGGERHYTITGSPVSTAL	300
<b>RESULT 5</b>					
PCB4_MOUSE	STANDARD; PRM;	403 AA.	Db	466 AQ 467	
ID	P57724;		Db	301 AQ 302	
AC			RP		
DT	16-OCT-2001 (Rel. 40, Created)		SEQUENCE FROM N.A.		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		STRAIN=CV_Columbia;		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		MEDLINE=20083488; PubMed=10617198;		
DE	POLY(rC) -binding protein 4 (Alpha-CP4).		RA		
GN	POBP4.		RA		
OS	Mus musculus (Mouse)		RA		
OC	Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;		RA		
OC	Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Mus.		RA		
OX	NCBI_TaxID=10090;		RA		
RN	[1]		RA		
RP	SEQUENCE FROM N.A.		RA		
RP	MEDLINE=20396135; PubMed=10936052;		RA		
RA	Markevey A. V.; Liebhäber S. A.;		RA		
RT	"Identification of two novel mammalian genes establishes a subfamily		RA		
RT	of KH-domain RNA-binding proteins.";		RA		
RL	Genomics 67:301-316(2000).		RA		
CC	-1- FUNCTION: SINGLE-STRANDED NUCLEIC ACID BINDING PROTEIN THAT		RA		
CC	BLINDS PREFERENTIALLY TO OLIGO DC (BY SIMILARITY).		RA		
CC	-1- SUBCELLULAR LOCATION: Nuclear (By similarity).		RA		
CC	-1- TISSUE SPECIFICITY: Ubiquitous (By similarity).		RA		
CC	-1- SIMILARITY: CONTAINS 3 KH DOMAINS.		RA		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		RA		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		RA		
CC	the European Bioinformatics Institute. There are no restrictions on its		RA		
CC	use by non-profit institutions as long as its content is in no way		RA		
CC	modified and this statement is not removed. Usage by and for commercial		RA		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		RA		
CC	or send an email to license@isb-sib.ch).		RA		
CC	DR EMBL; AF176328; AAC09239.1; - .		RA		
DR	MGI; MGI:1890471; Pcbp4.		RA		
DR	InterPro; IPR004087; KH_dom.		RA		
DR	InterPro; IPR004088; KH_type_1.		RA		
DR	Pfam; PF00013; KH-domain_3.		RA		
DR	SMART; SM00322; KH_3.		RA		
DR	PROSITE; PS50084; KH_TYPE_1; 3.		RA		
KW	Nuclear protein; RNA-binding; Ribonucleoprotein; DNA-binding;		RA		
KW	Repeat.		RA		
FT	DOMAIN 17 67 KH 1.		RA		
FT	DOMAIN 101 154 KH 2.		RA		
FT	DOMAIN 101 241 293 KH 3.		RA		
SQ	SEQUENCE 4 03 AA; 41398 MW; D4DBFFF490CEFE90A CRC64;		RA		
Query Match	7.28; Score 212.5; DB 1; Length 403;		RA		
Best Local Similarity	26.28; Pred. No. 2.9e-06;		RA		
Matches	79; Conservative 57; Mismatches 119; Indels 47; Gaps 12;		RA		
Qy	196 LPLRLVYPTQFGATIGKEGATIRNTMKTQSKIDYHRKENAGAAEKSTIRLSPGTGA 255		RA		

A	Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J., Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D., Minx P., Bentley D., Fulton B., Greco T., Kamp K., Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L., Nelson J., Spieth J., Ryan E., Andrews S., Geissel C., Layman D., Du H., Ali J., Berdoff A., Jones K., Drone M., Cotton J., Joshi C., Antonio B., Zidanic M., Strong C., Sun H., Lazar B., Yordan C., Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R., Swaby I. R., O'Shaughnessy A., Rodriguez M., Hoffman J., Tilli S., Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A., Chen E., Marra M., Martienssen R., McCombie W. R.;	Db	337 TREAGSCTI----EVNDSRTKGCGDECUTIVATESPDMKSMAVEAVLILQYEDND 391
A	"Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana."; Nature 402:769-777(1999). [2]	Db	337 TREAGSCTI----EVNDSRTKGCGDECUTIVATESPDMKSMAVEAVLILQYEDND 391
IN	CONCEPTUAL TRANSLATION.	Db	337 TREAGSCTI----EVNDSRTKGCGDECUTIVATESPDMKSMAVEAVLILQYEDND 391
A	Kieselbach T. Unpublished observations (JUL 2001).	Db	337 TREAGSCTI----EVNDSRTKGCGDECUTIVATESPDMKSMAVEAVLILQYEDND 391
IC	-1- SUBCELLULAR LOCATION: Nuclear (potential). -1- SIMILARITY: CONTAINS 5 KH DOMAINS.	Db	337 TREAGSCTI----EVNDSRTKGCGDECUTIVATESPDMKSMAVEAVLILQYEDND 391
IC	-1- CAUTION: REF 1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A GENE PREDICTION ERROR AND THE FUSION OF THE CDS FOR THIS PROTEIN WITH THE CDS FOR CHLOROPLAST PROTEASE RHOA.	Db	337 TREAGSCTI----EVNDSRTKGCGDECUTIVATESPDMKSMAVEAVLILQYEDND 391
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).	Db	337 TREAGSCTI----EVNDSRTKGCGDECUTIVATESPDMKSMAVEAVLILQYEDND 391
CC	EMBL; AL021710; CAA16717..1; ALT_SEQ EMBL; AL161548; CAB78839..1; ALT_SEQ DR InterPro; IPR004087; KH_domain_5. DR InterPro; IPR004088; KH_type_1..1. PFAM; PF00013; KH-domain_5. SMART; SM00322; KH_5. PROSITE; PS0008B; KH_TYPE_1..5. Hypothetical Protein; Nuclear protein; RNA-binding; Repeat.	Db	337 TREAGSCTI----EVNDSRTKGCGDECUTIVATESPDMKSMAVEAVLILQYEDND 391
CC	FT DOMAIN 35 99 KH 1. FT DOMAIN 138 210 KH 2. FT DOMAIN 311 380 KH 3. FT DOMAIN 394 455 KH 4. FT DOMAIN 535 599 KH 5. SEQUENCE 606 AA; 65760 MW; 61F135BBB8647C0C CRC64;	Db	337 TREAGSCTI----EVNDSRTKGCGDECUTIVATESPDMKSMAVEAVLILQYEDND 391
CC	Query Match Score 7.1%; Best Local Similarity 19.6%; Matches 89; Conservative 86; Mismatches 187; Indels 93; Gaps 14;	Db	337 TREAGSCTI----EVNDSRTKGCGDECUTIVATESPDMKSMAVEAVLILQYEDND 391
CC	QY 199 RLIVQTQPVGATIGKEATINRITKQTKSDVLRVKENAGAEKSITL----- 247	Db	337 TREAGSCTI----EVNDSRTKGCGDECUTIVATESPDMKSMAVEAVLILQYEDND 391
CC	QY 39 RILCPDVGVGVGSKGKVNAIRHTKAKTVFQDQLH-GSSQRVITYCSVKEKQEEIG 97	Db	337 TREAGSCTI----EVNDSRTKGCGDECUTIVATESPDMKSMAVEAVLILQYEDND 391
CC	QY 248 ---STPGETSAACKSILEIM-----HKAQDQIKFTEELPKTLAHNNFVRGLIGE 295	Db	337 TREAGSCTI----EVNDSRTKGCGDECUTIVATESPDMKSMAVEAVLILQYEDND 391
CC	QY 98 FTKSENEPILCICAADALLKVDIAVASDEBNNTKTNVDRDDKKECRLLVPFQSOSLIGKA 157	Db	337 TREAGSCTI----EVNDSRTKGCGDECUTIVATESPDMKSMAVEAVLILQYEDND 391
CC	QY 296 GRNLKKIBODDTKI----TISPLQETLTNPERTITVKGNVETCAK--AEEETMKK 347	Db	337 TREAGSCTI----EVNDSRTKGCGDECUTIVATESPDMKSMAVEAVLILQYEDND 391
CC	QY 158 GENIKRKRTRASVKVSKYDPSHVCAMYEDDNVVVSGSPESVKQALAVSA MYK 217	Db	337 TREAGSCTI----EVNDSRTKGCGDECUTIVATESPDMKSMAVEAVLILQYEDND 391
CC	QY 348 --RESYENDIASNHL-----QAHL-----IPLGLNLNALGL 375	Db	337 TREAGSCTI----EVNDSRTKGCGDECUTIVATESPDMKSMAVEAVLILQYEDND 391
CC	QY 218 NPRENITPLDTSQDVPAASVTPSDLNSNVYPTGFYSNQDHQGAGVPSY-FNALSV 276	Db	337 TREAGSCTI----EVNDSRTKGCGDECUTIVATESPDMKSMAVEAVLILQYEDND 391
CC	QY 376 ----FPPTSGMPPTSGPSPANTPPQFQESETETVHQFIPALSYGAIIGKQGOHIKO 430	Db	337 TREAGSCTI----EVNDSRTKGCGDECUTIVATESPDMKSMAVEAVLILQYEDND 391
CC	QY 277 SDFOGYAETAANPVFAASSLPVTHFGGSSRLVFKVLCPLCNMVRKGGSVTKR 336	Db	337 TREAGSCTI----EVNDSRTKGCGDECUTIVATESPDMKSMAVEAVLILQYEDND 391
CC	QY 431 LSRPAGASTKIAPEAPDAKVR----MVIITGGPEAQFRAQRGTYKIKEBNFVSPKE 484	Db	337 TREAGSCTI----EVNDSRTKGCGDECUTIVATESPDMKSMAVEAVLILQYEDND 391



RA	Kugler S., Grunweller A., Probst C., Klinger M., Muller P.K., Kruse C.,	"Vigilin contains a functional nuclear localisation sequence and is present in both the cytoplasm and the nucleus.";	Db	425 EGMRVLDINRMDYEVNINTDQHKFHRHL-----TICKSGANINRNDQKYVSRRP--
RT			Qy	527 QTPDENDQVVVKITGHFYACQVAQRKIQBILTOVKOHQ 564
RL			Db	473 -PDSEKSNLIRIEGDPQYQAKRELLBLASRMENER 508
CC	-!- FUNCTION: APPEARS TO PLAY A ROLE IN CELL STEROL METABOLISM.			
CC	IT MAY FUNCTION TO PROTECT CELLS FROM OVER-ACCUMULATION OF CHOLESTEROL.			
CC	-!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.			
CC	-!- SIMILARITY: CONTAINS 14 KH DOMAINS.			
CC	-!- SIMILARITY: STRONG, TO C.ELEGANS COB89.2.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions. This statement is as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).			
CC	EMBL: M64098; AAU35962.1; -.			
DR	GeneW; HGNC:4857; HDLRP.			
DR	MIM:142695; -.			
DR	InterPro: IPR004087; KH_dom.			
DR	InterPro: IPR004088; KH_type_1.			
DR	PFam: PF00013; KH-domain; 14.			
DR	SMART: SM00322; KH; 14.			
DR	PROSITE: PS00084; KH_TYPE_1; 14.			
KW	HDL; Lipid transport; Cholesterol metabolism; RNA-binding; Repeat; Phosphorylation; Nuclear protein.			
KW	Phosphorylation; Nuclear protein.			
FT	DOMAIN 158 229 KH 1.			
FT	DOMAIN 230 302 KH 2.			
FT	DOMAIN 303 371 KH 3.			
FT	DOMAIN 372 442 KH 4.			
FT	DOMAIN 443 514 KH 5.			
FT	DOMAIN 515 588 KH 6.			
FT	DOMAIN 589 660 KH 7.			
FT	DOMAIN 661 734 KH 8.			
FT	DOMAIN 735 807 KH 9.			
FT	DOMAIN 808 880 KH 10.			
FT	DOMAIN 881 979 KH 11.			
FT	DOMAIN 980 1059 KH 12.			
FT	DOMAIN 1060 1134 KH 13.			
FT	DOMAIN 1135 1209 KH 14.			
FT	MOD_RES 295 295 PHOSPHORYLATION (POTENTIAL).			
FT	MOD_RES 296 296 PHOSPHORYLATION (POTENTIAL).			
SQ	SEQUENCE 1268 AA; 141439 MW; 34ECB83D13A431 CRC64;			
Query Match	7.0%	Score 207.5; DB 1; Length 1268;		
Best Local Similarity	23.9%	Pred. No. 2.9e-05;		
Matches	95; Conservative	Mismatches 69; Indels 153; Gaps 18;		
Qy	202 VPTQFGAIKIGKEGATIRNTKQTOKIDVTKENAGAAKSITLSPGTTSACKSIL 261			
Db	157 IPEKHHRFVIGNGKLQDLELTATKIQIQLPQDD -- PSENQIKITGTKEIGEARHEVL 213			
Qy	262 EIMHKAQDKEFTEPLKTLAHNNFVG -- RLLICREGNRLKIEQDFTKTTISPLQEL 318			
Db	214 LI - SAEQDKRAVERLEVEAFHFPPTAGPNRLVG --- EIMQETQRTRINPP --- 260 *			
Qy	319 TLYNPERT-ITVKGNVETCAKAAEEIMKKIRESYNDIASMNL ---- QAHLIGLNLNA 372			
Db	261 --PSVNRTEIVFTGEKEQLAQAVARI-KK1YEKKKTIIIAVEVKKSQHVKYVGPCKNS 317			
Qy	373 LGLFPTSG --- MPPTS --- -GPF --- SAMTIPYQPFQESETETV --- HQ 410			
Db	318 LQEILLERTGVSVEIPSDSETVLLRGFPEKLGQALTEVAKANSFTVSSAAPSWLHR 377			
Qy	411 EIPALPSVGAITGKQOHIKOLSRFAGASIMAPRAPADKVRMVTITGPEAOFAKGRI 470			
Db	378 F-----LIGKKQNLAKITO-QMPKTHIEFFEGED --- KITLEGPTEDVNAQEQI 424			
Qy	471 YGKIKE-----ENFSVSPKEEYKLEAHIRVPSFAAGRIVGKGKTVNELQNLSSAEVVVRD 526			

Dbb	129	MIVRALLEERGNEDNGEDI---EISINLILIPHLMGCIIGKGRSRLREIEDLSAAKLF A 184
Qy	313	SPOLETLYNPERTITYKGNVETCAKEEIMKKIRESYNDIASMLNQAHLLPGLNLNA 372
Dbb	185	SPNO---LLSNDRLITINGVPDAIHATFYISLT-----LAFQME SPQKVNKR 231
Qy	373	LGLFPPTSGMPPPTSGPPSAMTPYQP-----EQSETEVY-----HOFIPALSGVALIG 422
Dbb	232	SIYQPT-----QNSVLDLSDQNTIFHQRNHOYHP - SDKLILSY 270
Qy	423	KQGHHIKQLS-----RFGASIKIAPAAEADAKV RAVVITGPQPAQFKAQGRIGYK 473
Dbb	271	KPNKNIPISSSTLSSMATQYTTSANATAFOQNFWPNVTVLGDYVISPAQGNHLLMMNF 330
Qy	474	IKEENFVSPEKEEVLAIRVPSFAAGRIVKGKTYVNEQNLNSSAEVVVPRDQTPDE-N 532
Dbb	331	VQDEFIDEK-----FVGNTIGDGHINSVKESTGCCIII---QDPVEGS 373
Qy	533	DQVYVKTGHTGHFYACQVA 549
Dbb	374	SERRATIRTEMASQOA 390

SULT 11	B2_HUMAN	STANDARD;	PRT;	365 AA.
	PCB2_HUMAN			
	Q15366;			
	30-MAY-2000 (Rel. 39, Created)			
	30-MAY-2000 (Rel. 39, Last sequence update)			
	16-OCT-2001 (Rel. 40, Last annotation update)			
	POLY(rC)-binding protein 2 (Alpha-CP2) (hnRNP-E2).			
	PCBP2.			
	Homo sapiens (Human).			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
	NCBI_TaxID=9606;			
	[1]			
	SEQUENCE FROM N.A.			
	MEDLINE=95331278; PubMed=7607214;			
	Leffers H., Degergaard K., Celis J.E.,			
	"Characterisation of two major cellular poly(rC)-binding human proteins, each containing three K-homologous (KH) domains."			
	Europ. J. Biochem. 230:447-453 (1995)			
-1-	FUNCTION: MAJOR CELLULAR POLY(rC)-BINDING PROTEIN. BINDS ALSO POLY(rU).			
-1-	SUBCELLULAR LOCATION: LOOSELY BOUND IN THE NUCLEUS. MAY SHUTTLE BETWEEN THE NUCLEUS AND THE CYTOPLASM.			
-1-	TISSUE SPECIFICITY: DETECTED IN ALL TISSUES EXAMINED.			
-1-	PIM: PHOSPHORYLATED. THE NON-PHOSPHORYLATED FORM(S) EXHIBITED STRONGEST POLY(rC)-BINDING ACTIVITY.			
	REFERENCES: 1. Leffers H., Degergaard K., Celis J.E., "Characterisation of two major cellular poly(rC)-binding human proteins, each containing three K-homologous (KH) domains." Europ. J. Biochem. 230:447-453 (1995). 2. Leffers H., Degergaard K., Celis J.E., "A cellular poly(rC)-binding protein with a KH domain." J. Cell. Physiol. 167:353-359 (1996).			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to licensee@isb-sib.ch).

EMBL: X78136; CAIA55015; - .  
HSSP: Q07244; 1KHM.  
Genew: HGNC:8648; PCBPF2.  
MIM: 601210; - .  
InterPro: IPR04087; KH\_dom.  
InterPro: IPR04088; KH\_type\_1.  
PFAM: PF00013; KH-domain; 3.  
SMART: SM00322; KH; 3.  
PROSITE: PS50084; KH\_TYPE\_1; 3.  
Nuclear protein; RNA-binding; Ribonucleoprotein; DNA-binding;  
phosphorylation; Repeat.  
DOMAIN: 13 75 KH 1.

FT	DOMAIN	97	162	KH 2.
FT	DOMAIN	287	351	KH 3.
SO	SEQUENCE	365 AA;	38580 MW;	43FE035D76FDC2C63 CRC64;
Query Match		6. 9%	Score 202. 5;	DB 1; Length 365;
Best Local Similarity	24. 0%	Pred NO. 1.1e-05;	Mismatches 119;	Indels 99; Gaps 13
Matches	87;	Conservative	58;	SEGNCPERITLLAGP-TNA 67
Qy	196 LPLRLVPTQFVGALIGKEGATIRNITQTKOSKIDVHRKENAGAAEKSITILSTPEGTSA 255			
Db	14 LTIRLIMHGKEVGSITGGKGESEVKRMREESGARINI			
Qy	256 ACKSLLIEHMKEAODIKFT			
Db	68 IFKAFAMIDKLEELISSIONTNSSTAASRPPTVLRAVVASOQGSLIGKGCKIKE TREST 127			
Qy	307 DTKITISPLQBLTYNPERRITYVG			
Db	128 GAQVQA-GDMLPNSTERATINGIPQSIIIEVKQICVVMLETLSQSPKGVTYPRPK 185			
Qy	348 -----RESYENDIASMNLOAHILPGLINLNALGLFPPTSGMP-----			
Db	186 PSSSPVIFAGQDRYSTGSDSASF--PHTPPSMCIN-----PDLEGPLEAYTIQCQYA 237			
Qy	384 -----PPTSG-----PFSAMMPPYPQEQQSETETVHQF'-PAL 415			
Db	238 IPQPDLTKLHOLAMQOSHFMTHNTGFGEISSLSPENYKGMWLDASQTSHELTIPND 297			
Qy	416 SVGALIGKGQGOHKIOLSRFAGASIKIA-PAEAAPDAKVRMVITGPPEAQFRGRTYKGKI 474			
Db	298 LIGCIGRQGAKINERIQMNSGAQIKIANPVEG--STDHQVITGSAASISLAQYLINVR L 355			
Qy	475 KEE 477			
Db	356 SSE 358			
RESULT 12				
PCB2_MOUSE				
ID	PCB2_MOUSE	STANDARD;	PRTR;	362 AA.
AC	Q611990; Q611383; Q62042;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Poly (rc)-binding protein 2 (Alpha-CP2) (Putative heterogeneous nuclear ribonucleoprotein X) (hnRNP X) (CBP) (CBP).			
GN	PCB2 OR HNRNPK OR HNRNPK			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TAXID=10090;			
RN	[1]	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).		
RP	STRAIN=C57BL/6 X 129G13	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).		
RC	STRAIN=C57BL/6 X 129G13	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).		
RX	MEDLINE=93376518; PubMed=8367306;			
RA	Hahn K.B., Kim G., Turich C., Smale S.T.;			
RT	"Isolation of a murine gene encoding a nucleic acid-binding protein with homology to hnRNP K."			
RL	Nucleic Acids Res. 21:3893-3894 (1993).			
RN	[2]	SEQUENCE FROM N.A. AND PARTIAL SEQUENCE (ISOFORM 2).		
RP	STRAIN=C57BL/6 X TISSUE=Liver;			
RC	STRAIN=C57BL/6 X CBA; TISSUE=Liver;			
RX	MEDLINE=94269912; PubMed=8208614;			
RA	Goller M., Funke B., Gehe-Becker C., Kroeger B., Lottspeich F., Horak I.;			
RT	"Murine protein which binds preferentially to oligo-C-rich single-stranded nucleic acids."			
RL	Nucleic Acids Res. 22:1895-1898 (1994).			
RN	[3]	SEQUENCE FROM N.A. (ISOFORM 3).		
RP	STRAIN=C57BL/6 X CBA; TISSUE=Liver;			
RC	STRAIN=C57BL/6 X CBA; TISSUE=Liver;			
RA	Horak I.;			
RL	Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[4]			

SEQUENCE FROM N.A. (ISOFORM 1).

RP VGLN\_CHICK STANDARD; PRT; 1270 AA.

RA Maleyev A.V.; Liephaber S.A.;

RT "Identificati<sup>n</sup> of two novel mammalian genes establishes a subfamily of KH-domain RNA-binding proteins.";

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: SINGLE-STRANDED NUCLEIC ACID BINDING PROTEIN THAT BINDS PREFERENTIALLY TO Oligo DC.

CC -!- SUBCELLULAR LOCATION: Nucleus.

CC -!- ALTERNATIVE PRODUCTS: ISOFORMS: 1 (SHOWN HERE), 2 AND 3; ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -!- SIMILARITY: CONTAINS 3 KH DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; L19661; AAA03105\_1; -;

CC DR EMBL; X97982; CAA66619\_1; -;

CC DR EMBL; AF236845; AAK14059\_1; -;

CC DR EMBL; AF236842; AAK14059\_1; JOINED.

CC DR EMBL; AF236843; AAK14059\_1; JOINED.

CC DR EMBL; AF236844; AAK14059\_1; JOINED.

CC DR HSSP; Q07244; 1KHM.

CC DR MGD; M31:108202; PCDP2.

CC DR InterPro; IPR004087; KH\_dom.

CC DR InterPro; IPR004088; KH\_type\_1.

CC DR Pfam; PF00013; KH\_domain; 3.

CC DR SMART; SM00322; KH; 3.

CC DR PROSITE; PS50034; KH\_TYPE\_1; 3.

CC KW Nuclear protein; RNA-binding; Ribonucleoprotein; DNA-binding;

CC KW Repeat; Alternative splicing.

FT DOMAIN 7 75 KH 1.

FT DOMAIN 13 97 162 KH 2.

FT DOMAIN 284 348 KH 3.

FT VARSPLIC 194 224 MISSING (IN ISOFORM 2).

FT VARSPLIC 263 275 MISSING (IN ISOFORM 3).

SQ SEQUENCE 362 AA; 38212 MW; 70CAF710E3F3C0 CR64;

Query Match 6.8%; Score 202; DB 1; Length 362;

Best Local Similarity 24.28%; Pred. No 1.1e-05;

Matches 87; Conservative 60; Mismatches 117; Indels 96; Gaps 14;

Db 196 LPLRLVYPTQFGAIGKEGATIRNITKOTOSKIDVHRRKENAGAEKSSTILSPCTGA 255

Qy 14 ITRRLMHGKEVGSITGKGESVKKREESGARINI----SEGNGPERITLLAGP--TNA 67

Qy 256 ACKSILEIMKHEAQDIKFET-----EEIPKLAINNFKVLRGKESRNUKKIEODT 306

Db 68 IFFAFMIIDKLEEDISSLSSMTNSTAASRPVTLVVAASOGSGSLIGKSOCKIKEIEST 127

Qy 307 DTKITISPLQEQLTLYNPERTIVKG---NVEETAKAEPEIMKKI----- 347

Db 128 GADYQVA-GMLPNSTERATTIAGIPOSIECYKQICVVMLESPPBKGTIPYRKPKSS 185

Qy 348 -----RESYENDIASMNIOQAHJPGINNLALGPPTSGMP----- 383

Db 186 PVIFAGGQDRYSTGDSASF--PHTPPSMCLN----PDLEGPLEATTIQQYAIPOP 237

Qy 384 -----PPIGGPP-SAMTPYPPQ-----PEQSEETVHQF-TPAISVG 418

Db 238 DLTKLHQALAMQSHFPMTHGNTGFS1ESSSPEVKGYAGLDASROTTSELTIPNDLG 297

Qy 419 AIGKQOQHKKQLSRFAGASIKIA-PAAEADAKYRVTITGPPEAFKAQGRIGKIKE 477

Db 298 CIIIGRQGAKINEIROMSGAQIKIANPVEG--STDROVITGSAASISLAQYLINVRSSE 355

Qy 182 QGSPGSV-SKQKPCDLPR-----LYPTOFVGAIGKEGATIRNITKOTOSKI-- 229

Db 483 EGDPQYQOARKELLEARMENERTKDLIEQFHRTIGDQGERIREKEPEVIN 542

Qy 230 ---DVHRKENAGAAEKSTTSSTPGTSACKSSTEINHKEQDKEFTEIPKTLAHNN 286

Db 543 FPPPAKSD-----IVQLQSPKNEVERCTKPMQKNAVDIVENSFSLSPPFQPHN 94

Qy 287 FVGRLIGKEGRNLKKEQDTKTTISPLQBLTNPRTITVGNVTCARAEETEMMK 346

Db 595 ---LICKGGANIKKIREFSNTKIDLPGRQAT---OROLQSQEQFTKLRLRHL-- 644

Qy 347 IRESYENDIASMNIOQAHJPGINNLALGPPTSGMP-----GLP----- PTSG ----- 381

Db	645	--TAQELANITEEVISLKSILNIGAKRFRIRSMECGVHTHPTGSGSATVT	701		RA	Zylebil M.J.;	
					RT	Cluster analysis of an extensive human breast cancer cell line	
					RT	Protein expression map database."	
QY	382	--MPPPTSGPPSAMTPPYQFQESETETY-		HQFIPAL-SVGAIIGKQGHIKQL	431	RT	Proteomics 2:212-223(2002).
DB	702	IRAQPTWRKPKRSCTWAEBKQTKSYTDLRAKPEYHKF-		--LIGGGGTTKRV	753	RL	[61]
QY	432	SRFAGASTKIAPEADAKYRMVLTGPPAEAQFKAQGRITYGKIKE-	ENIFVSKPEKEVLEA	490	RC	STRUCTURE BY NMR OF	
DB	754	RDTNTGARLTFPSEDDQEE-	LITMGTEEAVEAKKEQKELALINLDNVVEDSMVVDPKH	811	RA	Tjandra N.; Levens D.; Tjandra N.	
QY	491	HIVRESFAARGRTGKGCKTNEYLNQSLSAEVVYPRDQTPEENDOYVVKITGFYACQAQ	550	RT	"High Precision solution structure of the C-terminal KH domain of		
DB	812	H-----RHEVIRSQLREADEGVYMLPVLPTVSGHQSDKRYTLK--	-GAKDCWEAAK	861	RT	heterogeneous nuclear ribonucleoprotein K, a c-myc transcription	
QY	551	RKIQPLILQVKHOHQKALOSGPPQ	575	RT	factor."		
DB	862	KRIQIILEDL--EAQVIECTIQ	883	RL	J. Mol. Biol. 289:949-962(1999).		
				CC	CC		
				CC	FUNCTION: ONE OF THE MAJOR PRE-MRNA-BINDING PROTEINS. BINDS		
				CC	TERMINOUSLY TO POLY(C) SEQUENCES, LIKELY TO PLAY A ROLE IN		
				CC	THE NUCLEAR METABOLISM OF hnRNAs, PARTICULARLY FOR PRE-mRNAs THAT		
				CC	CONTAIN CYTIDINE-RICH SEQUENCES. CAN ALSO BIND POLY(C) SINGLE-		
				CC	STRANDED DNA.		
				CC	CC		
				CC	SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR; NUCLEOPLASM.		
				CC	CC		
				CC	ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE), AND 2; ARE		
				CC	PRODUCED BY ALTERNATIVE SPLICING.		
				CC	CC		
				CC	METHOD=MAFDI.		
				CC	CC		
				CC	MASS SPECTROMETRY: MW=50976.25;		
				CC	CC		
				CC	PHOSPHORYLATE.		
				CC	CC		
				CC	SIMILARITY: CONTAINS 3 KH DOMAINS.		
				CC	CC		
				CC	SWISS-PROT entry is copyright. It is produced through a collaboration		
				CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation		
				CC	the European Bioinformatics Institute. There are no restrictions on its		
				CC	use by non-profit institutions as long as its content is in no way		
				CC	modified and this statement is not removed. Usage by and for commercial		
				CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
				CC	or send an email to license@isb-sib.ch).		
				CC	CC		
				CC	EMBL; S74678; AAB20770_1;		
				DR	DR		
				CC	X72727; CAA51267_1;		
				DR	DR		
				CC	BC000355; AAH00355_1;		
				DR	DR		
				CC	B011980; AAH14980_1;		
				DR	DR		
				CC	D17711; BAA0566_1;		
				DR	DR		
				CC	PIR; AA2058; AA2058;		
				DR	DR		
				CC	S41495; S41495.		
				DR	PDB; 1KHM; 1Z-1AN-00.		
				DR	DR		
				CC	TRANSFAC; T02853;		
				DR	DR		
				CC	SWISS-2DPAGE; Q07244; HUMAN.		
				DR	DR		
				CC	SMART; SM00322; KH_3.		
				DR	DR		
				CC	PROSITE; P55004; KH_TYPE-1;		
				DR	DR		
				CC	Nuclear protein; RNA-binding; Repeat; DNA-binding;		
				DR	KW		
				CC	Phosphorylation; Alternative Splicing; 3D-structure.		
				FT	KH DOMAIN 42		
				FT	InterPro; IPR004087; KH dom.		
				FT	DR		
				FT	InterPro; IPR004088; KH_type_1.		
				FT	FT		
				FT	DOMAIN 144 197		
				FT	RNA-BINDING (RGG-BOX).		
				FT	FT		
				FT	DOMAIN 236 273		
				FT	KH 3		
				FT	387 439		
				FT	FT		
				FT	DOMAIN 54 421		
				FT	2 X 22 AA APPROXIMATE REPEATS.		
				FT	FT		
				FT	DOMAIN 54 54		
				FT	REPEAT 54 76		
				FT	1-1.		
				FT	REPEAT 399 421		
				FT	1-2.		
				FT	DOMAIN 245 329		
				FT	2 X 6 AA APPROXIMATE REPEATS.		
				FT	FT		
				FT	REPEAT 245 250		
				FT	2-1.		
				FT	REPEAT 324 329		
				FT	2-2.		
				FT	REPEAT 404 407		
				FT	5 X 4 AA REPEATS OF G-X-G-G.		
				FT	FT		
				FT	DOMAIN 289 294		
				FT	POLY-PRO.		
				FT	310 315		
				FT	VARSPLIC 459 463		
				FT	SGKPF -> ADVEGF (IN ISOFORM 2).		
				FT	CONFLICT 32 32		
				SO	A -> D (IN REF.).		
				SEQUENCE	463 AA; 5076 MW; QFT7DE169B24064A CRC64;		

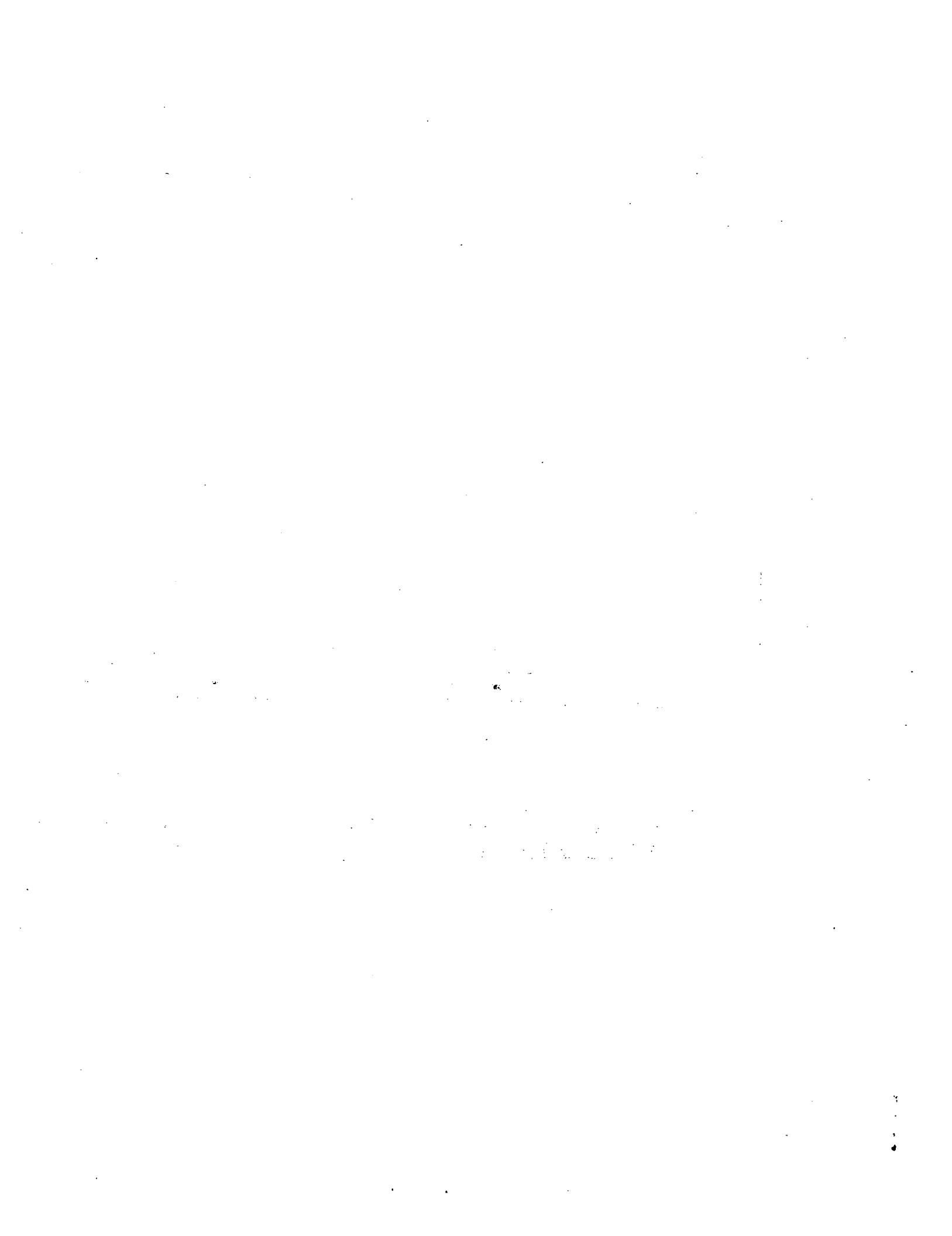
Query Match	6.4%	Score 190; DB 1; Length 463;	DR HSSP; Q07244; 1KH <sub>M</sub> , InterPro; IPI004087; KH_dom.
Best Local Similarity	20.7%	Pred. No. 8.9e-05; 54; Mismatches 131; .Indels 148; Gaps 13;	DR Pfam; PF00013; KH_domain; 3.
Matches	87	Conservative	DR SMART; SM00322; KH; PROSITE; PS50084; KH_TYPE_1; 3.
Oy	198	LRLVPTQFVAAIIGKAGATIRNITKOTQSKIDVHRKENAGAAEKSITLSTPEGTSAAC 257. ♫	DR Nuclear protein; RNA-binding; Ribonucleoprotein; Repeat; DNA-binding;
Db	45	LRILQSKNAGAVIGGGKNNIKALRDYNASVSV --PDSSGPERILSISADIETIGETL 101.	KW phosphorylation.
Qy	258	KSILEIMHKEAQ--DIKFTEIP-----LKLIAHNNFVGRLIG 293	KW DOMAIN 4.2 9/2 KH 1.
Db	102	KKLIPTLEGQOLPSPTATSQLPLESDAVECLNYQHYKGSDDFCEURLIHQSLAGGLIG 161	FT DOMAIN 144 19/7 KH 2.
Qy	294	KEGRNKKIEQDIDTDTKTTISPLQELTNPERITVKGNVETCAKEEEIMKKTRES--- 350	FT DOMAIN 236 27/3 RNA-BINDING (RGG-BOX).
Db	162	VRGAKIKEELRLENQT-TIKLQECCPHSTDRVLLGGKPPRVEIKILIDLISESPIK 219	FT DOMAIN 387 43/9 KH 3.
Qy	351	-----YENDIASMNLQIAHLIPGLNLNALGLF-- -PPTSG-- -MPP--- 384	FT DOMAIN 54 4/21 2 X 22 AA APPROXIMATE REPEATS.
Db	220	GRAQPYDPNFYDETYDGGFTMMEDDRGRPVGPMRGRGGFDRMPPGRGMPGRMPPSRRI 279	FT DOMAIN 76 1/1 1-1.
Qy	385	-----PTSGPPS-----AMTPPPYPQ-----F 400	FT REPEAT 399 4/21 1-2.
Db	280	YDDMSPRRGPPPPPGRGGSSARNLPLPPPPRGGLMDAYDRGRPGDRYDMGVGF 339	FT REPEAT 245 3/29 2 X 6 AA REPEATS OF D-R-R-G-R-P.
Qy	401	EQSET-----ETVHQFIPALSY 417	FT REPEAT 250 2/2.
Db	340	SABETWSAIDWPSWSEPMQAYEPQGGSQDYSYAGGRGSYGDGLGPIITITQVTIPKDLA 399	FT REPEAT 324 3/29 2-2.
Qy	418	GALIKQOQHIIKOLSRFAGASIKI-APAEAPDAVKRMVILTGPPAQFKAQGRYGIKE 476	FT REPEAT 59 4/07 5 X 4 AA REPEATS OF G-X-G-G.
Db	400	GSIIGKQGRIKORIHRHESGASTKIDELEG-- -SEDRTITGTQDQ1QNAQYLQNSVKQ 457	FT DOMAIN 59 6/2 POLY-PRO.
Qy			FT DOMAIN 308 3/15 POLY-PRO.
Qy			SQ SEQUENCE 463 AA; 50960 MW; A6A92E645EFB1B6 CRC64;
			Query Match 6.4%; Score 190; DB 1; Length 463;
			Best Local Similarity 20.7%; Pred. No. 8.9e-05;
			Matches 87; Conservative 54; Mismatches 131; Indels 148; Gaps 13;
Qy	198	LRLVPTQFVAAIIGKAGATIRNITKOTQSKIDVHRKENAGAAEKSITLSTPEGTSAAC 257	Db 45 LRLIQSKNAGAVIGGGKNNIKALRDYNASVSV --POSSPERILSISADIETIGIRL 101
Qy	258	KSILEIMHKEAQ--DIKFTEIP-----LKLIAHNNFVGRLIG 293	Db 102 KKIPIPLEEGQLQPSFTATSQLPLESDAVECLNYQHFKGSDDFCEURLIHOSSLAGLIG 161
Qy	258	KSILEIMHKEAQ--DIKFTEIP-----LKLIAHNNFVGRLIG 293	Db 102 KKIPIPLEEGQLQPSFTATSQLPLESDAVECLNYQHFKGSDDFCEURLIHOSSLAGLIG 161
Qy	294	KEGRNKKIEQDIDTDTKTTISPLQELTNPERITVKGNVETCAKEEEIMKKTRES--- 350	Db 162 VRGAKIKEELRLENQT-TIKLQECCPHSTDRVLLGGKPPDRVVECIKILDISESPIK 219
Qy	351	-----YENDIASMNLQIAHLIPGLNLNALGLF-- -PPTSG-- -MPP--- 384	Db 220 GRAQPYDPNFYDETYDGGFTMMEDDRGRPVGPMRGRGGFDRMPPGRGMPGRMPPSRRI 279
Qy	385	-----PTSGPPS-----AMTPPPYPQ-----F 400	Db 280 YDDMSPRRGPPPPPGRGGSSARNLPLPPPPRGGLMDAYDRGRPGDRYDMGVGF 339
Qy			Db 400 GSIIKGQGRIKORIHRHESGASIKIDFLEG-- -SEDRTITGTQDQ1QNAQYLQNSVKQ 457

## RESULT 15

ROK\_RABBIT  
 ID ROK\_RABBIT STANDARD; PRT; 463 AA.  
 AC 01949;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Heterogeneous nuclear ribonucleoprotein K (hnRNP K).  
 GN hnRNP\_K.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagidae.  
 RN [1] PTNSQPPS-  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=99207122; PubMed=10101190-  
 RA Thiele B.J., Berger M., Huth A., Reimann I., Schwarz K., Thiele H.;  
 RT "Tissue-specific translational regulation of alternative rabbit  
 RT 15'-1-poxigenase mRNAs differing in their 3'-untranslated regions.";  
 RL Nucleic Acids Res. 27:1829-1836 (1999).  
 CC - FUNCTION: ONE OF THE MAJOR PRE-mRNA-BINDING PROTEINS. BINDS  
 CC TENTACULARLY TO POLY(C) SEQUENCES. LIKELY TO PLAY A ROLE IN THE  
 CC NUCLEAR METABOLISM OF hnRNAs, PARTICULARLY FOR PRE-mRNAs THAT  
 CC CONTAIN CYTIDINE-RICH SEQUENCES. CAN ALSO BIND POLY(C) SINGLE-  
 CC STRANDED DNA.  
 CC - SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR; NUCLEOPLASM (BY  
 CC SIMILARITY).  
 CC - PTM: PHOSPHORYLATED (BY SIMILARITY).  
 CC - SIMILARITY: CONTAINS 3 KH DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/amrnone/>  
 CC or send an email to license@isb-sib.ch).  
 DR EMBL; AJ003204; CAA05815.1; -.

Search completed: April 24, 2003, 16:13:28  
 Job time : 16 secs



GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 24, 2003, 16:11:07 ; Search time 40 Seconds

(without alignments)  
1928.803 Million cell updates/sec

Title: US-09-897-778-176

Perfect score: 2956 Sequence: MRLYIGNLSENAAPSDLES . . . . . VRQHQQQKALQSPPQSRRK 579

Scoring table: BLOSUM62 Gap open 10.0 , capext 0.5

Searched: 908170 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Benesed\_101002:\*

1: /SID52/gcldata/geneseqp/geneseqp-emb1/AA1980 DAT:\*

2: /SID52/gcldata/geneseqp/geneseqp-emb1/AA1981.DAT:\*

3: /SID52/gcldata/geneseqp/geneseqp-emb1/AA1982.DAT:\*

4: /SID52/gcldata/geneseqp/geneseqp-emb1/AA1983 DAT:\*

5: /SID52/gcldata/geneseqp/geneseqp-emb1/AA1984.DAT:\*

6: /SID52/gcldata/geneseqp/geneseqp-emb1/AA1985.DAT:\*

7: /SID52/gcldata/geneseqp/geneseqp-emb1/AA1986 DAT:\*

8: /SID52/gcldata/geneseqp/geneseqp-emb1/AA1987.DAT:\*

9: /SID52/gcldata/geneseqp/geneseqp-emb1/AA1988.DAT:\*

10: /SID52/gcldata/geneseqp/geneseqp-emb1/AA1989.DAT:\*

11: /SID52/gcldata/geneseqp/geneseqp-emb1/AA1990 DAT:\*

12: /SID52/gcldata/geneseqp/geneseqp-emb1/AA1991.DAT:\*

13: /SID52/gcldata/geneseqp/geneseqp-emb1/AA1992.DAT:\*

14: /SID52/gcldata/geneseqp/geneseqp-emb1/AA1993.DAT:\*

15: /SID52/gcldata/geneseqp/geneseqp-emb1/AA1994.DAT:\*

16: /SID52/gcldata/geneseqp/geneseqp-emb1/AA1995.DAT:\*

17: /SID52/gcldata/geneseqp/geneseqp-emb1/AA1996 DAT:\*

18: /SID52/gcldata/geneseqp/geneseqp-emb1/AA1997.DAT:\*

19: /SID52/gcldata/geneseqp/geneseqp-emb1/AA1998.DAT:\*

20: /SID52/gcldata/geneseqp/geneseqp-emb1/AA1999 DAT:\*

21: /SID52/gcldata/geneseqp/geneseqp-emb1/AA2000.DAT:\*

22: /SID52/gcldata/geneseqp/geneseqp-emb1/AA2001.DAT:\*

23: /SID52/gcldata/geneseqp/geneseqp-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	2956	100.0	579	AAB11328
2	2956	100.0	579	ABB74960
3	2943	99.6	579	AAB11365
4	2943	99.6	579	ABB74997
5	2943	99.6	579	ABB75053
6	2943	99.6	579	ABB75054
7	2938	99.4	586	ABB75048
8	2814	95.2	619	ABG21963
9	2733	92.5	583	ABG12592
10	2190	74.1	577	AYX30649

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	2956	100.0	579	AAB11328
2	2956	100.0	579	ABB74960
3	2943	99.6	579	AAB11365
4	2943	99.6	579	ABB74997
5	2943	99.6	579	ABB75053
6	2943	99.6	579	ABB75054
7	2938	99.4	586	ABB75048
8	2814	95.2	619	ABG21963
9	2733	92.5	583	ABG12592
10	2190	74.1	577	AYX30649

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	2956	100.0	579	AAB11328
2	2956	100.0	579	ABB74960
3	2943	99.6	579	AAB11365
4	2943	99.6	579	ABB74997
5	2943	99.6	579	ABB75053
6	2943	99.6	579	ABB75054
7	2938	99.4	586	ABB75048
8	2814	95.2	619	ABG21963
9	2733	92.5	583	ABG12592
10	2190	74.1	577	AYX30649

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match Length	DB ID	Description
1	2956	100.0	579	AAB11328
2	2956	100.0	579	ABB74960
3	2943	99.6	579	AAB11365
4	2943	99.6	579	ABB74997
5	2943	99.6	579	ABB75053
6	2943	99.6	579	ABB75054
7	2938	99.4	586	ABB75048
8	2814	95.2	619	ABG21963
9	2733	92.5	583	ABG12592
10	2190	74.1	577	AYX30649

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match Length	DB ID	Description
1	2956	100.0	579	AAB11328
2	2956	100.0	579	ABB74960
3	2943	99.6	579	AAB11365
4	2943	99.6	579	ABB74997
5	2943	99.6	579	ABB75053
6	2943	99.6	579	ABB75054
7	2938	99.4	586	ABB75048
8	2814	95.2	619	ABG21963
9	2733	92.5	583	ABG12592
10	2190	74.1	577	AYX30649

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match Length	DB ID	Description
1	2956	100.0	579	AAB11328
2	2956	100.0	579	ABB74960
3	2943	99.6	579	AAB11365
4	2943	99.6	579	ABB74997
5	2943	99.6	579	ABB75053
6	2943	99.6	579	ABB75054
7	2938	99.4	586	ABB75048
8	2814	95.2	619	ABG21963
9	2733	92.5	583	ABG12592
10	2190	74.1	577	AYX30649

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match Length	DB ID	Description
1	2956	100.0	579	AAB11328
2	2956	100.0	579	ABB74960
3	2943	99.6	579	AAB11365
4	2943	99.6	579	ABB74997
5	2943	99.6	579	ABB75053
6	2943	99.6	579	ABB75054
7	2938	99.4	586	ABB75048
8	2814	95.2	619	ABG21963
9	2733	92.5	583	ABG12592
10	2190	74.1	577	AYX30649

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match Length	DB ID	Description
1	2956	100.0	579	AAB11328
2	2956	100.0	579	ABB74960
3	2943	99.6	579	AAB11365
4	2943	99.6	579	ABB74997
5	2943	99.6	579	ABB75053
6	2943	99.6	579	ABB75054
7	2938	99.4	586	ABB75048
8	2814	95.2	619	ABG21963
9	2733	92.5	583	ABG12592
10	2190	74.1	577	AYX30649

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match Length	DB ID	Description
1	2956	100.0	579	AAB11328
2	2956	100.0	579	ABB74960
3	2943	99.6	579	AAB11365
4	2943	99.6	579	ABB74997
5	2943	99.6	579	ABB75053
6	2943	99.6	579	ABB75054
7	2938	99.4	586	ABB75048
8	2814	95.2	619	ABG21963
9	2733	92.5	583	ABG12592
10	2190	74.1	577	AYX30649

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match Length	DB ID	Description
1	2956	100.0	579	AAB11328
2	2956	100.0	579	ABB74960
3	2943	99.6	579	AAB11365
4	2943	99.6	579	ABB74997
5	2943	99.6	579	ABB75053
6	2943	99.6	579	ABB75054
7	2938	99.4	586	ABB75048
8	2814	95.2	619	ABG21963
9	2733	92.5	583	ABG12592
10	2190	74.1	577	AYX30649

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match Length	DB ID	Description
1	2956	100.0	579	AAB11328
2	2956	100.0	579	ABB74960
3	2943	99.6	579	AAB11365
4	2943	99.6	579	ABB74997
5	2943	99.6	579	ABB75053
6	2943	99.6	579	ABB75

XX Isolated polypeptide comprising an immunogenic portion of a lung tumor  
PT protein is used for detecting and monitoring progression of lung cancer  
PT in a patient.

XX Claim 3; Page 186-188; 261pp; English.

XX This invention describes a novel isolated polypeptide (I) which  
CC comprises an immunogenic portion of a lung tumor protein or variant (P2)  
CC which have cytostatic activity. The polypeptides and polynucleotides are  
CC used in compositions and vaccines to inhibit the development of cancer,  
CC especially lung cancer, in a patient. Methods described in the invention  
CC can be used to monitor the progression of a cancer by carrying out the  
CC detection at subsequent time points and comparing the results from the  
CC different time points. CD4+ and/or CD8+ T-cells isolated from a patient  
CC are treated with P2, polynucleotides encoding P2 or antigen presenting  
CC cells expressing P2 and then administered to the patient to inhibit  
CC development of cancer.

XX Sequence 579 AA;

Query Match 100.0%; Score 2956; DB 21; Length 579;  
Best Local Similarity 100.0%; Pred. No. 9.7e-268;  
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNKLYIGNSENAPSDELSIFKDAPISGPPFLVKTYGAFVDCPDESWALKATEALSGK 60  
Db 1 MNKLYIGNSENAPSDELSIFKDAPISGPPFLVKTYGAFVDCPDESWALKATEALSGK 60

Qy 61 IELHGKPIEVHSPVKRQIRKQIQRNTPHQLWEVDLSDLLYQGYVESECVQNTDSETA 120  
Db 61 IELHGKPIEVHSPVKRQIRKQIQRNTPHQLWEVDLSDLLYQGYVESECVQNTDSETA 120

Qy 121 VVNTYSSKDOQARQALDKINGFOLENTFLKVAYIPDMAQNPQPRGRGLGORGSS 180  
Db 121 VVNTYSSKDOQARQALDKINGFOLENTFLKVAYIPDMAQNPQPRGRGLGORGSS 180

Qy 181 RQSPGSYSVKOKCDPLPLRLLPTQFGAIGKEGATIRNITKQTOSKIDVHRKENAGAA 240  
Db 181 RQSPGSYSVKOKCDPLPLRLLPTQFGAIGKEGATIRNITKQTOSKIDVHRKENAGAA 240

Qy 241 EKSITLSTPEGTSAACKSILEIMHKBAQDIKFTEEIPKLTLAHNNFVGRLIGKEGRNLK 300  
Db 241 EKSITLSTPEGTSAACKSILEIMHKBAQDIKFTEEIPKLTLAHNNFVGRLIGKEGRNLK 300

Qy 301 KIEQDTDTKITISPLQELTLYNPERTITYKGIVETCAKAEEFIMKKIRESYENDIASMNL 360  
Db 301 KIEQDTDTKITISPLQELTLYNPERTITYKGIVETCAKAEEFIMKKIRESYENDIASMNL 360

Qy 361 QAHLPGNLNAAGLFPLPTSGMPPTSGPPSAMTTPYQFQEQSETETVHQFIPALSGA 420  
Db 361 QAHLPGNLNAAGLFPLPTSGMPPTSGPPSAMTTPYQFQEQSETETVHQFIPALSGA 420

Qy 421 IGKOGQHIKQLSRFAGASIKTAAPEADAKYRMVITGPPEAQFQGRGYKIKEENFY 480  
Db 421 IGKOGQHIKQLSRFAGASIKTAAPEADAKYRMVITGPPEAQFQGRGYKIKEENFY 480

Qy 481 SPKEEVKLEAHIRVPSAAGRVIGKGGKTVNELQNISSAEVVPRDQTPENDQVVKIT 540  
Db 481 SPKEEVKLEAHIRVPSAAGRVIGKGGKTVNELQNISSAEVVPRDQTPENDQVVKIT 540

Qy 541 GHFYACQVAQRKIQEILTQVKOHOOQKALQSGPPSRK 579  
Db 541 GHFYACQVAQRKIQEILTQVKOHOOQKALQSGPPSRK 579

RESULT 2  
ID ABB74960 standard; Protein; 579 AA.  
XX AC ABB74960;  
DT 01-MAY-2002 (first entry)

XX Human lung tumour L523S protein sequence SEQ ID NO:176.  
DE XX Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;  
KW KW immune response.  
XX OS Homo sapiens.  
XX PN WO200200174-A2.  
XX PD 03-JAN-2002.  
XX PP 28-JUN-2001; 2001WO-US21065.  
XX PR 28-JUN-2000; 2000US-0606421.  
PR 02-AUG-2000; 2000US-0630940.  
PR 15-AUG-2000; 2000US-0643597.  
PR 09-OCT-2000; 2000US-0662286.  
PR 09-DEC-2000; 2000US-0685696.  
PR 02-DEC-2000; 2000US-0735705.  
PR 07-MAY-2001; 2001US-0850716.  
XX PA (CORIXA CORP.  
XX PI Wang T, Wang A, Skeiki YAW, Li SX, Kalos MD, Henderson RA;  
PI McNeill PD, Fanger N, Rettner MW, Marnerakis M,  
PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;  
XX DR N PSDB; AB149119.  
WPI: 2002-090513/12.  
XX PT Polynucleotides encoding lung tumor polypeptides, useful for treating  
PT lung cancer or stimulating an immune response -  
XX PS Example 2; Page 267-268: 374pp; English.  
XX CC The present invention describes human lung tumour proteins. Human lung  
CC tumour proteins and polynucleotides have cytostatic and immunostimulant  
CC activities and can be used in vaccine production. Compositions  
CC comprising the lung tumour proteins, polynucleotides, antibodies,  
CC fusion proteins, T cell populations, or antigen presenting cells that  
CC express the lung tumour proteins are useful for treating lung cancer or  
CC stimulating an immune response. ABL48939 to ABL4930 and ABB7496 to  
CC ABB75070 represent sequences used in the exemplification of the present  
CC invention.  
XX SQ Sequence 579 AA;  
Query Match 100.0%; Score 2956; DB 23; Length 579;  
Best Local Similarity 100.0%; Pred. No. 9.7e-246;  
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MNKLYIGNSENAPSDELSIFKDAPISGPPFLVKTYGAFVDCPDESWALKATEALSGK 60  
Db 1 MNKLYIGNSENAPSDELSIFKDAPISGPPFLVKTYGAFVDCPDESWALKATEALSGK 60  
Qy 61 IELHGKPIEVHSPVKRQIRKQIQRNTPHQLWEVDLSDLLYQGYVESECVQNTDSETA 120  
Db 61 IELHGKPIEVHSPVKRQIRKQIQRNTPHQLWEVDLSDLLYQGYVESECVQNTDSETA 120  
Qy 121 VVNTYSSKDOQARQALDKINGFOLENTFLKVAYIPDMAQNPQPRGRGLGORGSS 180  
Db 121 VVNTYSSKDOQARQALDKINGFOLENTFLKVAYIPDMAQNPQPRGRGLGORGSS 180  
Qy 181 RQSPGSYSVKOKCDPLPLRLLPTQFGAIGKEGATIRNITKQTOSKIDVHRKENAGAA 240  
Db 181 RQSPGSYSVKOKCDPLPLRLLPTQFGAIGKEGATIRNITKQTOSKIDVHRKENAGAA 240  
Qy 241 EKSITLSTPEGTSAACKSILEIMHKBAQDIKFTEEIPKLTLAHNNFVGRLIGKEGRNLK 300  
Db 241 EKSITLSTPEGTSAACKSILEIMHKBAQDIKFTEEIPKLTLAHNNFVGRLIGKEGRNLK 300  
Qy 301 KIEQDTDTKITISPLQELTLYNPERTITYKGIVETCAKAEEFIMKKIRESYENDIASMNL 360  
Db 301 KIEQDTDTKITISPLQELTLYNPERTITYKGIVETCAKAEEFIMKKIRESYENDIASMNL 360  
Qy 361 QAHLPGNLNAAGLFPLPTSGMPPTSGPPSAMTTPYQFQEQSETETVHQFIPALSGA 420  
Db 361 QAHLPGNLNAAGLFPLPTSGMPPTSGPPSAMTTPYQFQEQSETETVHQFIPALSGA 420  
Qy 421 IGKOGQHIKQLSRFAGASIKTAAPEADAKYRMVITGPPEAQFQGRGYKIKEENFY 480  
Db 421 IGKOGQHIKQLSRFAGASIKTAAPEADAKYRMVITGPPEAQFQGRGYKIKEENFY 480  
Qy 481 SPKEEVKLEAHIRVPSAAGRVIGKGGKTVNELQNISSAEVVPRDQTPENDQVVKIT 540  
Db 481 SPKEEVKLEAHIRVPSAAGRVIGKGGKTVNELQNISSAEVVPRDQTPENDQVVKIT 540  
Qy 541 GHFYACQVAQRKIQEILTQVKOHOOQKALQSGPPSRK 579  
Db 541 GHFYACQVAQRKIQEILTQVKOHOOQKALQSGPPSRK 579

Db 301 KIEQDTDTKTTISPLQELTNPERTITYRGNVETCAKEEIMRKIRESYENDIASMNL 360  
 QY 361 QAHLIGLNALNGLFPPTSGMPPTSGPAMPPYQEPOSETETVHOFIPALSGAI 420  
 Db 361 QAHLIGLNALNGLFPPTSGMPPTSGPAMPPYQEPOSETETVHOFIPALSGAI 420  
 QY 421 IKGQGHIKQLSRFAGASIKIATAPADAKVRMVTITGPEAOPKAQGRIGKIKEENFV 480  
 Db 421 IKGQGHIKQLSRFAGASIKIATAPADAKVRMVTITGPEAOPKAQGRIGKIKEENFV 480  
 QY 481 SPKEEVKLEAHIRVPSFAAGRVIGGGKTVELNQLSSAVVVPRDQTPDENQVVKKLT 540  
 Db 481 SPKEEVKLEAHIRVPSFAAGRVIGGGKTVELNQLSSAVVVPRDQTPDENQVVKKLT 540  
 QY 541 GHFYACOVAQRKIQEITLQYQHQOOKALSGPOSRRK 579.  
 Db 541 GHFYACOVAQRKIQEITLQYQHQOOKALSGPOSRRK 579

**RESULT 3**

AAB11365 standard; Protein; 579 AA  
 ID AAB11365 standard; Protein; 579 AA  
 XX  
 AC  
 XX  
 DT 21-FEB-2001 (first entry)  
 DE Human lung cancer associated antigen L523S.  
 XX  
 KW Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;  
 KW vaccine; detection.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2000061612-A2.  
 XX  
 PD 19-OCT-2000.  
 XX  
 PF 03-APR-2000; 20000WO-US08896.  
 XX  
 PR 02-APR-1999; 99US-0285479.  
 PR 17-DEC-1999; 99US-0466396.  
 PR 30-DEC-1999; 99US-0476196.  
 PR 10-JAN-2000; 2000US-0480884.  
 PR 22-FEB-2000; 2000US-0510376.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Wang T, Fan L;  
 XX  
 PS Claim 3; Page 259-261; 261pp; English.  
 XX  
 DR WPI; 2000-G28399/60.  
 DR N-PSDB; AAC66035.

CC This invention describes a novel isolated polypeptide (I) which  
 CC comprising an immunogenic portion of a lung tumor protein or variant (P2)  
 CC which have cytostatic activity. The polypeptides and polynucleotides are  
 CC used in compositions and vaccines to inhibit the development of cancer,  
 CC especially lung cancer, in a patient. Methods described in the invention  
 CC can be used to monitor the progression of a cancer by carrying out the  
 CC detection at subsequent time points and comparing the results from the  
 CC different time points. CD4+ and/or CD8+ T-Cells isolated from a patient  
 CC are treated with P2, Polynucleotides encoding P2 or antigen presenting  
 CC cells expressing P2 and then administered to the patient to inhibit  
 CC development of cancer.  
 XX Sequence 579 AA;

Query Match 99.6%; Score 2943; DB 21; Length 579;  
 Best Local Similarity 99.7%; Pred. No. 1, 3e-24;  
 Matches 577; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MNKLYIGNLSENAAPSDELSIFKDAKIPVSGPFLVKTGAYAFVDCPDESWALKAEALSGK 60  
 Db 1 MNKLYIGNLSENAAPSDELSIFKDAKIPVSGPFLVKTGAYAFVDCPDESWALKAEALSGK 60  
 QY 61 IELHGKPIEVEHSVPKORITKLQTINIPPHQWEVLDSLLQYGTVESCOVNNTSETA 120  
 Db 61 IELHGKPIEVEHSVPKORITKLQTINIPPHQWEVLDSLLQYGTVESCOVNNTSETA 120  
 QY 121 VNNTYSSKDOARQALDKLNGFOLENTFLTKVAYIPDEMAAQONPLOQPRGRGLGORGSS 180  
 Db 121 VNNTYSSKDOARQALDKLNGFOLENTFLTKVAYIPDEMAAQONPLOQPRGRGLGORGSS 180  
 QY 181 ROGSPGSYSKQKPCDLPLRLYPTQFGAIGKEGATIRNITKOTOSKIDYHRKENAGAA 240  
 Db 181 ROGSPGSYSKQKPCDLPLRLYPTQFGAIGKEGATIRNITKOTOSKIDYHRKENAGAA 240  
 QY 241 EKSITLSTPEGTSAAACKSILEIMHKEAQDIKTFEEPLKILAHNNFGRGLIGKERNLK 300  
 Db 241 EKSITLSTPEGTSAAACKSILEIMHKEAQDIKTFEEPLKILAHNNFGRGLIGKERNLK 300  
 QY 301 KIEODDTDTKTTISPLQELTLYNPERTITVKGNVETCAKEEIMKKRESYENDIASMNL 360  
 Db 301 KIEODDTDTKTTISPLQELTLYNPERTITVKGNVETCAKEEIMKKRESYENDIASMNL 360  
 QY 361 QAHLIPGLNLNALGLFPPTSGMPPTSGPSSAMTIPPQFEQSETETVHQFIPALSGAI 420  
 Db 361 QAHLIPGLNLNALGLFPPTSGMPPTSGPSSAMTIPPQFEQSETETVHLFIPALSGAI 420  
 QY 421 IGKOGQHKQLSFGAGASIKIAAEAPDAKVRMVTITGPPEAQFKAGRIGKIKENFV 480  
 Db 421 IGKOGQHKQLSFGAGASIKIAAEAPDAKVRMVTITGPPEAQFKAGRIGKIKENFV 480  
 QY 481 SPKEEVKLEAHIRVPSFAAGRIVGKGKTVNELQNLSSAEEVYVPRDQTPDENQVVKIT 540  
 Db 481 SPKEEVKLEAHIRVPSFAAGRIVGKGKTVNELQNLSSAEEVYVPRDQTPDENQVVKIT 540  
 QY 541 GHFYACOVAQRKIQEITLQYQHQOQQKALSGQPQSRRK 579  
 Db 541 GHFYACOVAQRKIQEITLQYQHQOQQKALSGQPQSRRK 579

**RESULT 4**

ABB74997  
 ID ABB74997 standard; Protein; 579 AA.  
 XX  
 AC ABB74997;  
 XX  
 DT 01-MAY-2002 (first entry)  
 XX  
 DE Human lung tumour L523S protein sequence SEQ ID NO:348.  
 XX  
 KW Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;  
 KW immune response.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2000174-A2.  
 XX  
 PR 03-JAN-2002.  
 XX  
 PR 28-JUN-2001; 2001WO-US21065.  
 XX  
 PR 28-JUN-2000; 2000US-0606421.  
 PR 02-AUG-2000; 2000US-060940.  
 PR 21-AUG-2000; 2000US-063597.  
 PR 15-SEP-2000; 2000US-0662786.  
 PR 09-OCT-2000; 2000US-0685696.  
 PR 12-DEC-2000; 2000US-0735705.

PR 07-MAY-2001; 2001US-0850716.  
 XX  
 PA (CORTI-) CORIXA CORP.  
 XX Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;  
 PI McNeill PD, Fanger N, Rettter MW, Marnerakis M, Fanger GR;  
 PI Vedick TS, Carter D, Watamabe Y, Peckham DW;  
 XX DR WPI; 2002-090513/12.  
 DR N-PSBB; ABL49254.  
 XX PT Polynucleotides encoding lung tumor polypeptides, useful for treating  
 PT lung cancer or stimulating an immune response -  
 XX PS Example 2; Page 330-332; 374PP; English.  
 XX CC The present invention describes human lung tumour proteins. Human lung  
 CC tumour proteins and polynucleotides have cytostatic and immunostimulant  
 CC activities, and can be used in vaccine production. Compositions  
 CC comprising the lung tumour proteins, polynucleotides, antibodies,  
 CC fusion proteins, T cell populations, or antigen presenting cells that  
 CC express the lung tumour proteins are useful for treating lung cancer or  
 CC stimulating an immune response. ABL48059 to ABL49300 and ABB74946 to  
 CC ABB5070 represent sequences used in the exemplification of the present  
 CC invention.  
 XX SQ Sequence 579 AA;

Query Match 99.6%; Score 2943; DB 23; Length 579;  
 Best Local Similarity 99.7%; Pred. No. 1.3e-244;  
 Matches 577; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MNKLYIGNNSENAAPSVDLSEIFKDAKIVSGPFLVKTGAYFVDCPDESWALKAEIALSGK 60  
 Db 1 MNKLYIGNNSENAAPSVDLSEIFKDAKIVSGPFLVKTGAYFVDCPDESWALKAEIALSGK 60  
 Qy 61 IELHGKPIEVEHSYSPKRQIRKQLQIRNIPPHLOWEVLDSLVQGVYSECVNTDSETA 120  
 Db 61 TEURGKPIEVEHSYSPKRQIRKQLQIRNIPPHLOWEVLDSLVQGVYSECVNTDSETA 120  
 Qy 121 VNVNTYSSKDQARQALDKINGFQLENFTLKVAYIPDEMAAQNPLOQPRGRGLGORGSS 180  
 Db 121 VNVNTYSSRDQARQALDKINGFQLENFTLKVAYIPDETAQNPLOQPRGRGLGORGSS 180  
 Qy 181 RQSGPGSVSKQPKCDPLPLRLYPTQFGAIGREGATRNTKTKOSKIDYHRKENAGAA 240  
 Db 181 RQSGPGSVSKQPKCDPLPLRLYPTQFGAIGREGATRNTKTKOSKIDYHRKENAGAA 240  
 Qy 241 EKSITILSPEGTSAAACKSILETMHEKAQDIKFTEELPLKILAHNNFVGRLIGKEGRNLK 300  
 Db 241 EKSITILSPEGTSAAACKSILETMHEKAQDIKFTEELPLKILAHNNFVGRLIGKEGRNLK 300  
 Qy 301 KIEODDTKIKTISPLQELTLYNPERTITYKGNETCAKEETMKKTRSYENDIASMNL 360  
 Db 301 KIEODDTKIKTISPLQELTLYNPERTITYKGNETCAKEETMKKTRSYENDIASMNL 360  
 Qy 361 QAHLIPGMLNAGLGFPTSGMPPTSGPSAMTPPYQFEGESETEVHQFIPALSGVA 420  
 Db 361 QAHLIPGMLNAGLGFPTSGMPPTSGPSAMTPPYQFEGESETEVHQFIPALSGVA 420  
 Qy 421 IGKOGHIIKQLSRFAGASTKIAKAAEPDAKVRMVIITGPPEAQFKAQGRIYGKIKEENFY 480  
 Db 421 IGKOGHIIKQLSRFAGASTKIAKAAEPDAKVRMVIITGPPEAQFKAQGRIYGKIKEENFY 480  
 Qy 481 SPREEVKLEAHIRVPSFAAGRVTGGGKTVNQNLSAAEVVPRDODPNDQVVKIT 540  
 Db 481 SPREEVKLEAHIRVPSFAAGRVTGGGKTVNQNLSAAEVVPRDODPNDQVVKIT 540  
 Qy 541 GHFYACQVAKRKIQEILITQVKHQOKQALSGPQSRK 579  
 Db 541 GHFYACQVAKRKIQEILITQVKHQOKQALSGPQSRK 579

RESULT 5  
 ABB75053 standard; Protein: 579 AA.  
 ID ABB75053;  
 XX AC ABB75053;  
 DT 01-MAY-2002 (first entry)  
 DE Human lung tumour L523S recombinant protein sequence SEQ ID NO 446  
 XX KW Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;  
 KW Immune response.  
 XX OS Homo sapiens.  
 XX PN WO200200174-A2.  
 XX PD 03-JAN-2002.  
 XX PF 28-JUN-2001; 2001WO-US21065.  
 XX PR 28-JUN-2000; 2000US-0606421.  
 PR 02-AUG-2000; 2000US-0630940.  
 PR 21-AUG-2000; 2000US-0643597.  
 PR 15-SEP-2000; 2000US-0662786.  
 PR 09-OCT-2000; 2000US-0685896.  
 PR 12-DEC-2000; 2000US-0735705.  
 PR 07-MAY-2001; 2001US-0850716.

XX PA (CORTI-) CORIXA CORP.  
 XX PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;  
 PI McNeill PD, Fanger N, Rettter MW, Marnerakis M, Fanger GR;  
 PI Vedick TS, Carter D, Watanabe Y, Peckham DW;  
 XX DR WPI; 2002-090513/12.  
 XX DR AB49297.  
 XX PT Polynucleotides encoding lung tumor polypeptides, useful for treating  
 PT lung cancer or stimulating an immune response -  
 XX PS Claim 2; Page 365-367; 374PP; English.  
 XX PS The present invention describes human lung tumour proteins. Human lung  
 CC tumour proteins and polynucleotides have cytostatic and immunostimulant  
 CC activities, and can be used in vaccine production. Compositions  
 CC comprising the lung tumour proteins, polynucleotides, antibodies,  
 CC fusion proteins, T cell populations or antigen presenting cells that  
 CC express the lung tumour proteins are useful for treating lung cancer or  
 CC stimulating an immune response. ABL48059 to ABL49300 and ABB74946 to  
 CC ABB5070 represent sequences used in the exemplification of the present  
 CC invention.  
 XX SQ Sequence 579 AA;

Query Match 99.6%; Score 2943; DB 23; Length 579;  
 Best Local Similarity 99.7%; Pred. No. 1.3e-244;  
 Matches 577; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MNKLYIGNNSENAAPSVDLSEIFKDAKIVSGPFLVKTGAYFVDCPDESWALKAEIALSGK 60  
 Db 1 MNKLYIGNNSENAAPSVDLSEIFKDAKIVSGPFLVKTGAYFVDCPDESWALKAEIALSGK 60  
 Qy 121 VNVNTYSSKDQARQALDKINGFQLENFTLKVAYIPDEMAAQNPLOQPRGRGLGORGSS 180  
 Db 121 VNVNTYSSRDQARQALDKINGFQLENFTLKVAYIPDETAQNPLOQPRGRGLGORGSS 180  
 Qy 181 RQSGPGSVSKQPKCDPLPLRLYPTQFGAIGREGATRNTKTKOSKIDYHRKENAGAA 240  
 Db 181 RQSGPGSVSKQPKCDPLPLRLYPTQFGAIGREGATRNTKTKOSKIDYHRKENAGAA 240  
 Qy 241 EKSITILSPEGTSAAACKSILETMHEKAQDIKFTEELPLKILAHNNFVGRLIGKEGRNLK 300  
 Db 241 EKSITILSPEGTSAAACKSILETMHEKAQDIKFTEELPLKILAHNNFVGRLIGKEGRNLK 300  
 Qy 301 KIEODDTKIKTISPLQELTLYNPERTITYKGNETCAKEETMKKTRSYENDIASMNL 360  
 Db 301 KIEODDTKIKTISPLQELTLYNPERTITYKGNETCAKEETMKKTRSYENDIASMNL 360  
 Qy 361 QAHLIPGMLNAGLGFPTSGMPPTSGPSAMTPPYQFEGESETEVHQFIPALSGVA 420  
 Db 361 QAHLIPGMLNAGLGFPTSGMPPTSGPSAMTPPYQFEGESETEVHQFIPALSGVA 420  
 Qy 421 IGKOGHIIKQLSRFAGASTKIAKAAEPDAKVRMVIITGPPEAQFKAQGRIYGKIKEENFY 480  
 Db 421 IGKOGHIIKQLSRFAGASTKIAKAAEPDAKVRMVIITGPPEAQFKAQGRIYGKIKEENFY 480  
 Qy 481 SPREEVKLEAHIRVPSFAAGRVTGGGKTVNQNLSAAEVVPRDODPNDQVVKIT 540  
 Db 481 SPREEVKLEAHIRVPSFAAGRVTGGGKTVNQNLSAAEVVPRDODPNDQVVKIT 540  
 Qy 541 GHFYACQVAKRKIQEILITQVKHQOKQALSGPQSRK 579  
 Db 541 GHFYACQVAKRKIQEILITQVKHQOKQALSGPQSRK 579

Qy 121 VNVNTYSSKDQARQALDKINGFQLENFTLKVAYIPDEMAAQNPLOQPRGRGLGORGSS 180  
 Db 121 VNVNTYSSRDQARQALDKINGFQLENFTLKVAYIPDETAQNPLOQPRGRGLGORGSS 180  
 Qy 181 RQSGPGSVSKQPKCDPLPLRLYPTQFGAIGREGATRNTKTKOSKIDYHRKENAGAA 240

Do	181	RQGSPGSVKOKPCDPLPLRLLVPTQFYGAIGKEGATIRNITKOTOSKIDVRKENAGAA	240	comprising the lung tumour proteins, polynucleotides, antibodies, fusion proteins, T cell populations, or antigen presenting cells that express the lung tumour proteins are useful for treating lung cancer or stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to ABB75048 represent sequences used in the exemplification of the present invention.
Qy	241	EKSITLSTPEGTSAACKSLEIMHKEADIKTFEIPKILAHNNFVGRLIGKEGRNLK	300	
Db	241	EKSITLSTPEGTSAACKSLEIMHKEADIKTFEIPKILAHNNFVGRLIGKEGRNLK	300	
Qy	301	KIEQDTDTKITSPLQELTLYNPERTITYKGNEYTCAKAEEETMKKIRESYENDIASMNL	360	
Db	301	KIEQDTDTKITSPLQELTLYNPERTITYKGNEYTCAKAEEETMKKIRESYENDIASMNL	360	
Qy	361	QAHLIPLGNLNALGLFPPTSGMPPTSAMTPYQFEQSETETVHQFIPALPSVGA	420	Query Match Score 99.6%; Best Local Similarity 99.7%; Pred. No. 1. 3e-244; Mismatches 0; Indels 0; Gaps 0;
Db	361	QAHLIPLGNLNALGLFPPTSGMPPTSAMTPYQFEQSETETVHQFIPALPSVGA	420	Sequence 579 AA;
Qy	421	IGKQGHIKKOLSRFGASAKTAPADEKVRAYITLGPEAQFKAGRIYKIKEENFV	480	Query Match Score 99.6%; Best Local Similarity 99.7%; Pred. No. 1. 3e-244; Mismatches 0; Indels 0; Gaps 0;
Db	421	IGKQGHIKKOLSRFGASAKTAPADEKVRAYITLGPEAQFKAGRIYKIKEENFV	480	Sequence 579 AA;
Qy	481	SPKEYKLEAHITRVSSFAAGRIVKGKTYNEQNOLNSSEAVVYPRDTPDENQYVVKIT	540	Query Match Score 99.6%; Best Local Similarity 99.7%; Pred. No. 1. 3e-244; Mismatches 0; Indels 0; Gaps 0;
Db	481	SPKEYKLEAHITRVSSFAAGRIVKGKTYNEQNOLNSSEAVVYPRDTPDENQYVVKIT	540	Sequence 579 AA;
Qy	541	GHFYACOVAQRKIQELTQYVQHOOQAKALOSGPQSRK	579	Query Match Score 99.6%; Best Local Similarity 99.7%; Pred. No. 1. 3e-244; Mismatches 0; Indels 0; Gaps 0;
Db	541	GHFYACOVAQRKIQELTQYVQHOOQAKALOSGPQSRK	579	Sequence 579 AA;
RESULT 6				
ID	ABB75054	standard; protein; 579 AA.		
XX				
AC	ABB75054;			
XX				
DT	01-MAY-2002	(first entry)		
XX				
DE	Human lung tumour L523S recombinant protein sequence SEQ ID NO:449.			
XX				
KW	Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine; immune response.			
XX				
OS	Homo sapiens.			
XX				
PN	WO200200174 A2.			
XX				
PD	03-JAN-2002.			
XX				
PF	28-JUN-2001; 2001WO-US21065.			
XX				
PR	28-JUN-2000; 2000US-0606421.			
PR	02-AUG-2000; 2000US-0630940.			
PR	15-SEP-2000; 2000US-0642786.			
PR	09-OCT-2000; 2000US-0685896.			
PR	12-DEC-2000; 2000US-073597.			
PR	07-MAY-2001; 2001US-0850716.			
XX				
PA	(CORI-) CORIXA CORP.			
XX				
PI	Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;			
PI	McNeill PD, Fanger N, Rettie MW, Marneris M, Fanger GR;			
PI	Vedick TS, Carter D, Watanabe Y, Peckham DW;			
XX				
DR	WPI: 2002-090513/12.			
DR	N-PSDB; ABL49299.			
XX				
PT	Polynucleotides encoding lung tumor polypeptides, useful for treating lung cancer or stimulating an immune response -			
XX				
PS	Claim 2; Page 368-369; 374pp; English.			
XX				
CC	The present invention describes human lung tumour proteins. Human lung tumour proteins and polynucleotides have cytostatic and immunostimulant activities, and can be used in vaccine production. Compositions			
CC				



QY	1	MNKLYGNLSENAASDLSIFDKAKTEVSGPEPLVKTGYAFVDCPDESMALKATEALUSKG	PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity
Db	37	MNKLYGNLSENAASDLSIFDKAKTEVSGPEPLVKTGYAFVDCPDESMALKAFSGK	PT
QY	61	IEHKGKLEVEHSVPKRQERKQLQIRNTPHQIWEVLDSLLVQGVESCEQVNTDSETA	XX
Db	97	IELHGKLEVEHSVPKRQERKQLQIRNTPHQIWEVLDSLLVQGVESCEQVNTDSETA	XX
QY	121	WVNNTYSSKDOARQA -- LDKLNGFQLENFTLVKAYIDPEMAAQONPQLOPQRGRRLGQ	CC
Db	157	WVNNTYSSKDOARQALDKLNGFQLENFTLVKAYIDPEMAAQONPQLOPQRGRRLGQ	CC
QY	177	RGSRQGSPSUSVKQPCDPLRLLPVPOFGAIIGKECATIRNTKQSKIDVHKEN	CC
Db	217	RGSRQGSPSUSVKQPCDPLRLLPVPOFGAIIGKECATIRNTKQSKIDVHKEN	CC
QY	237	AGRAEKSTTILSTPBTSAACKSTLEIHKAEQDIKTEELPLKLAHNNVGRLICKEG	CC
Db	277	AGRAEKSTTILSTPBTSAACKSTLEIHKAEQDIKTEELPLKLAHNNVGRLICKEG	CC
QY	297	RNLKKIQDQDTKTTSPBTSAACKSTLEIHKAEQDIKTEELPLKLAHNNVGRLICKEG	CC
Db	337	RNLKKIQDQDTKTTSPBTSAACKSTLEIHKAEQDIKTEELPLKLAHNNVGRLICKEG	CC
QY	357	SMLQAHLIPGLNLNALGLFRPTSGMPPTSGPSSAMTPYPOEQSETETVQFIALS	CC
Db	397	SMLQAHLIPGLNLNALGLFRPTSGMPPTSGPSSAMTPYPOEQSETETVQFIALS	CC
QY	417	VGAITIGKQGHIKOLSRFAGASKIAPIAEPDAKVRMVTGPPAAQFGRIYGKKE	CC
Db	457	VRALISROGDHIKOLSRFAGASSKIAPIAEPDAKVRMVTGPPAAQFGRIYGKKE	CC
QY	477	ENVSPREEVKLEAHRVPAFAACRVIGRKTVNELQNLSSAEVVVPDQTDENDOV	CC
Db	517	ENVSPREEVKLEAHRVPAFAACRVIGRKTVNELQNLSSAEVVVPDQTDENDOV	CC
QY	537	VKTGHYVAQOVAQRIKQETLTVOKHQDQOKALOSGPQRK	CC
Db	577	VKTGHYVAQOVAQRIKXELTVQHQDQOKALOSGPQRK	CC
RESULT 9			
ID	ABG12592	standard; Protein: 583 AA.	
XX	AC	ABG12592;	
XX	DT	18-FEB-2002 (first entry)	
XX	DE	Novel human diagnostic protein #12583.	
XX	KW	Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.	
XX	OS	Homo sapiens.	
XX	PN	WO200175067-A2.	
XX	PD	11-OCT-2001-	
XX	PF	30-MAR-2001; 2001WO-US08631.	
XX	PR	31-MAR-2000; 2000US-0540317.	
XX	PR	23-AUG-2000; 2000US-0649167.	
XX	PA	(HYSE-) HYSEQ INC.	
XX	PI	Liu C, Tang YT;	
XX	DR	WPI; 2001-639352/73.	
DR	N-PSDB; AAS6779.		
XX			
QY	357	SMNLOAHLPGLNALGLFPPTSGMPPTSGPSSAMTPYPOEQSETETVQFIALS	16
Db	361	SMNLOAHLPGLNALGLFPPTSGMPPTSGPSSAMTPYPOEQSETETVQFIALS	20
QY	237	AGAAKSITLSPGTGSAKSILEIMKEADQIKPEETPLKLAHNHNVGRLICKEG	296
Db	241	AGAAKSITLSPGTGSAKSILEIMKEADQIKPEETPLKLAHNHNVGRLICKEG	300
QY	297	RNLKTEQDTDTKITSQELTYNPERTITYVGNYETCAKAEETMKIRESYENDIA	356
Do	301	RNLKTEQDTDTKITSQELTYNPERTITYVGNYETCAKAEETMKIRESYENDIA	360
QY	357	SMNLOAHLPGLNALGLFPPTSGMPPTSGPSSAMTPYPOEQSETETVQFIALS	416
Db	361	SMNLOAHLPGLNALGLFPPTSGMPPTSGPSSAMTPYPOEQSETETVQFIALS	420
QY	417	VGAITIGKQGHIKOLSRFAGASKIAPIAEPDAKVRMVTGPPAAQFGRIYGKKE	476
Db	421	VRALISROGDHIKOLSRFAGASSKIAPIAEPDAKVRMVTGPPAAQFGRIYGKKE	490
QY	477	ENVSPREEVKLEAHRVPAFAACRVIGRKTVNELQNLSSAEVVVPDNDQV	536
Db	481	ENVSPREEVKLEAHRVPAFAACRVIGRKTVNELQNLSSAEVVVPDNDQV	540
QY	537	VKITGHYACQVAQRIKQETLTVOKHQDQOKALOSGPQRK	579

Db	541	VKIGHFFYACQVAQRKIQEILTQVKQHOOQKALQSQQPSRKK	583
RESULT 1.0			
ID	AAT30649	AAT30649 standard; Protein; 577 AA.	
XX	AAT30649;		
XX	17-NOV-1999	(first entry)	
XX	A murine c-myc coding region determinant binding protein.		
XX	c-myc coding region determinant binding protein; CRD-BP; tumor; c-myc; endonucleolytic attack; half-life; breast cancer; colon cancer; pancreatic cancer.		
XX	Mus musculus.		
XX	W0946524-A2.		
XX	PD 16-SEP-1999.		
XX	PF 05-MAR-1999;	99WO-US04897.	
XX	PR 09-MAR-1998;	98US-0077372.	
XX	PA (WISCONSIN ALUMNI RES FOUND.		
XX	PI Ross J;		
XX	PT WPI; 1999-551506/46.		
XX	DR N-PDB; AAZ101617.		
XX	Diagnosing presence or absence of a tumor in a human by examining c-myc coding region determinant-binding protein		
XX	Example: Fig 1A-D; 79pp; English.		
XX	The present sequence represents a murine c-myc coding region determinant binding protein (CRD-BP). The presence or absence of a tumor can be determined by determining the levels of CRD-BP present in the suspect tissue, where the CRD-BP shields c-myc RNA from endonucleolytic attack and so prolongs its half-life. The methods are used for diagnosing presence or absence of a tumor in a human, especially breast, colon and pancreatic cancer. They are also used to inhibit cancer cell growth.		
XX	Sequence 577 AA;		
XX	Query Match 74.1%; Score 2190; DB 20; Length 577;		
XX	Best Local Similarity 74.1%; Pred. No. 1e-79; Mismatches 62; Indels 16; Gaps 8;		
XX	Matches 434; Conservative 1; Score 2190; DB 20; Length 577;		
XX	1 MNKLYIGNLSENRAPSDELESFKDAKIPYSGPFLVTKGAYFVDCPDSWALKATEALSGK 60		
XX	1 MNKLYIGNLSENVTPADLEYFAEKHISTSGQFLVKSAYAFVDCPDBHWAHKATEALSGK 60		
XX	61 VVNTYSSDQAQDQLNLGPOLENPLKTVYLPDEMAAQNQLQOPRGR->GPENGGGGFSRGK 120		
XX	61 VELQGKRLEMENSVPKQRSKQIQRNIPPOLRNEVLSLLAQYTGVENCEQYNTSETA 120		
XX	1.21 VVNTYSSDQAQDQLNLGPOLENPLKTVYLPDEMAAQNQLQOPRGR->GPENGGGGFSRGK 179		
XX	1.21 VVNTYSSDQAQDQLNLGPOLENPLKTVYLPDEMAAQNQLQOPRGR->GPENGGGGFSRGK 176		
Db	Qy 180 SROGSP--GSVSKQPCDPLRLVPTQVAGTKECATINITQTSKIDVIRKEN 236		
Db	Qy 177 PROSPGVAAAGAPAKQPVDPRLVPTQVAGTKECATINITQTSKIDVIRKEN 236		
Db	Qy 237 AGAAEKSTTSLSPGTTSAACKSLSLEIMMKBAQD1KTFTEPIPLKILAHNFVGRLLKG 296		
Db	Qy 237 AGAAEKSTTSLSPGTTSAACKSLSLEIMMKBAQD1KTFTEPIPLKILAHNFVGRLLKG 296		

CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AB00010-ABc30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at [ftp.wipo.int/pub/published.pct\\_sequences](http://wipo.int/pub/published.pct_sequences).

XX	Sequence	614 AA;	Score 66.0%;	Length 614;	
Query Match	Best Local Similarity	Score 1950.5; Pred. No. 4.8e-159;	DB 22;	Gaps 8;	
Matches	386;	Mismatches 95;	Indels 29;		
Qy	1 MNKLYIGNLSENAAPSDELSIEFKDAKTAPIVPSGPFLVKTYGAFVDFCPDESMAALKATEALSSR	60;			
Db	17 MNKLYIGNLPAVTADDLRLQEGDKKLPLKGVLKSGYAFVDFPDQNWAIRATELSSK	76			
Qy	61 IELHGKPIEVEHHSVPKRQRIRKLQLTRNIPHLQNEYLDSLLVVOYGVVESEQEYNTDSETA	120			
Db	77 VELHGKIMEDVYSVSKLRSKIQIERNIPHLQWELDGLLAQQTGTVNEEQVNIDTEA	136			
Qy	121 VVNVTYSSKPOARQALDKLNGFQLENFTLVAYIPDEMAAQONPLQQPRQRGLGORGSS	180			
Qy	137 VVNVTYATREAKIAKEKLSSHOHENYFSKISYDVEVSSPSPQ--RRQR--GDHSSR	192			
Db	181 RQG-SPGSVSKOPCDPLRLVPTQFGATIGKESATINITKOTOSKTDVHKKENAGA	239			
Qy	193 EGSHAPGTQSQRQIDEPLRILVPTQFGATIGKESLTINKNITKOTOSRDIHKENSGA	252			
Qy	240 AKSITITLSPPEGTSAACKSLTEIMKHEADIKFTETIPPLKTLAHNFVGRNLIGKEGRNL	299			
Db	253 AEPVPTTHAEPSTSEACRMLLEIMOREADETKLAEPIPLKTLAHNLGVGLIGKEGRNL	312			
Qy	300 KKIEQDITDKITISPLQELTLYNPERTIVKGNTCAKAREEIMKKIRESYENDTASMN	359			
Db	313 KKLEHECTKTRITSSLDLTYNPERTIVKGTVEAASAEIMEKKLRAEFDMLAVN	372			
Qy	360 LQAHLIPGNNLNALGLF----PPTS--GMPPPT----SGPPSAMTP----P 396				
Db	373 QQANLIPGNNLNALGIFTGSLVSPAPGAPPAPPAHYHPFTTHEGYFSSLYPHQQFQGP	432			
Qy	397 YPOFEQ-SPEIETYHQTFPLPSVGAITKGQHOKLSRFAAGSIKAPAAEPAADAKYRMVI	455			
Db	433 FPHHSYDPEIVNLFPTQAGAIIKGKAHKOLAFAGASIKIAPAAEPAVDSEMV	492			
Qy	456 ITGPPEAOFKAOGRIGYKIKEENFVSPKEVKLEAHIRVPSFAAGRVIGKGGRTVNELQN	515			
Db	493 ITGPPEAOFKAOGRIGFHKIKEENFFNPKKEVKLEAHIRVPSSTAGRVIGKGGKTVNELQN	552			
Qy	516 LSSEAIVVYERDQPDDENDOQVYKIGTGHYACQVAQRTQELUTQVKHQOQ	566			
Db	553 LTSAEVIVRDQTPDENEEVIVRLIGHFESQTAQRKIREIVQQVKQOEK	603			
RESULT 12					
XX	AAU16163 Human novel secreted protein, Seq ID 1116.				
ID	AAU16163 standard; Protein; 620 AA.				
XX	AAU16163;				
AC					
XX	07-NOV-2001 (first entry)				
Dt					
XX					
DE					
XX					
Human; immunosuppressive; antiarthritic; antirheumatic;					
KW cytostatic; cardiotonic; vasotonic; cerebroprotective; nootropic;					
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;					
KW vulnerary; secreted protein; rheumatoid arthritis; cardiovascular disorder; cardiac arrest;					
KW hyperproliferative disorder; cardiovascular disorder; cerebral ischaemia; angiogenesis;					
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;					

PR 21-SEP-2000; 2000US-02342243.

PR 21-SEP-2000; 2000US-0234274.

PR 21-SEP-2000; 2000US-0234274.

KW corneal infection; wound healing; epithelial cell proliferation;  
 KW skin ageing; food additive; preservative; antiproliferative.

KW Homo sapiens.

XX WO200155322-A2.

XX 02-AUG-2001.

PD 02-AUG-2001; 2001WO-US01341.

XX PR 17-JAN-2001; 2001WO-US01341.

XX PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-020515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 18-AUG-2000; 2000US-0225270.

PR 22-AUG-2000; 2000US-0226651.

PR 22-AUG-2000; 2000US-0226658.

PR 22-AUG-2000; 2000US-0227182.

PR 14-AUG-2000; 2000US-0225158.

PR 14-AUG-2000; 2000US-0225279.

PR 18-AUG-2000; 2000US-0226229.

PR 22-AUG-2000; 2000US-0226651.

PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229387.

PR 01-SEP-2000; 2000US-0229313.

PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229559.

PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 14-SEP-2000; 2000US-0231413.

PR 14-SEP-2000; 2000US-0232399.

PR 08-SEP-2000; 2000US-0232400.

PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0233063.

PR 14-SEP-2000; 2000US-0233064.

PR 14-SEP-2000; 2000US-0233065.

PR 21-SEP-2000; 2000US-0234223.

PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235184.  
 PR 27-SEP-2000; 2000US-0235334.  
 PR 27-SEP-2000; 2000US-0235336.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 02-OCT-2000; 2000US-0238802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-023935.  
 PR 13-OCT-2000; 2000US-023935.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 01-NOV-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0241617.  
 PR 08-NOV-2000; 2000US-0241674.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0244476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0245523.  
 PR 08-NOV-2000; 2000US-0241524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0245527.  
 PR 08-NOV-2000; 2000US-0241528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0244610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 17-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0244207.  
 PR 17-NOV-2000; 2000US-0244208.  
 PR 17-NOV-2000; 2000US-0244210.  
 PR 17-NOV-2000; 2000US-0244210.  
 PR 17-NOV-2000; 2000US-02419211.  
 PR 17-NOV-2000; 2000US-02419212.  
 PR 17-NOV-2000; 2000US-0244213.  
 PR 17-NOV-2000; 2000US-0244214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249298.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0251160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251886.  
 PR 05-DEC-2000; 2000US-0251887.  
 PR 05-DEC-2000; 2000US-0251888.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0234097.

PR 05-JAN-2001; 2001US-0259678.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Barash SC, Ruben SM;

XX DR N-PSDB; AAS26150.

XX PT New nucleic acid molecules encoding 461 human secreted proteins for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives -

XX PS Claim 11; SEQ ID NO 1116; 980pp; English.

XX CC The invention relates to isolated nucleic acid molecules and their encoded secreted proteins. The nucleic acids and proteins are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Antibodies to the proteins can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms, of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection, and many other disorders listed in the specification. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein, vitamins, minerals, cofactors and other nutritional components. The present sequence represents a novel secreted protein of the invention.

Query Match Score 1946.5; DB 22; Length 620;

Best Local Similarity 65.1%; Pred. No. 1..1-e-186;

Matches 385; Conservative 81; Mismatches 96; Indels 29; Gaps 8;

Qy 1 MNKLYTGNLSENAAPSDLSFISIFKDAFKPSVSGPFLVKYGAFFVDCPDDESWALKAEALSGK 60

Db 23 MNKLVIGNLSPAVTADLQLFEDRKPLAGQVLLSAGYATVDPQNWAIRATELSK 82

Qy 61 IELHGKPIEYHSVPKRORIKLQIQRNIPPHLQWYDLSLUVQGVVESQEYNTDSETA 120

Db 83 VELHGKIMEVDYSVSKLRSRKLQIQRNIPPHLQWEVDLGIAQXGTVENQEYNTDETA 142

Qy 121 VVNVTYSSKQDARQDALDKNGFQLENFTLKVAYIPDEMAAQNPQLQPRGRGLGQRSS 180

Db 143 VVNVTYATBREAKAMEKLSHOPENNSFKISKYIDPDEVSSPSPPQ-RAQF-GDHSSR 198

Qy 181 RQG-SPGSYSKQKPCDLPRLILYPTQFGVAGLKGEATINNIKTQSKADVHRKENAGA 239

Db 199 EQGHAPGTSQARQDLPRLILYPTQFGVAGLKGEATINNIKTQSKADVHRKENAGA 258

Qy 240 AEKSITLILSTPPEGTSACKSLIBIMHQAQIKFTEPIPKLAIHNFYGLIGKEGNRL 299

Db 259 AEKPVTHATPPEGTSEACRMILETMQREADENKLAEPIPLKILARINGLYGLIGKEGNRL 318

Qy 300 KTKEDDTDTITISPLQDFTLYMPERTITVGNVENCRAABEEITMKTRSYENDTAMN 359

Db 319 KTKHEHTGKTISSQLDLSYINPERTTVKGTVEACASAIEIMKKLAEFENDMLAVN 378

Qy 360 LQAHLTGPNLNALGFL-----PPTS--GMPPPT-----SGPSAMTP----P 396

Db 379 QANLJPGNLSAIGFSTGLSLSPLSPAGPRAAPPAPYHFTPSLPHQGP 438

Qy 397 YPOFQQ -SETETVHQFIPALPSGATIGKQGQHIKQLSRFGASIKIAPAEAPDAKVRVVI 455

PR biodversity PS Claim 20; SEQ ID NO 37154; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and poly peptide (II) sequences, (I) is useful as hybridisation probes, and in recombinant production of (II). The poly nucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (III). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. BG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published\\_pct\\_sequences](ftp://wipo.int/pub/published_pct_sequences).

Query Match	Score	Length
Best Local Similarity	63.2%	DB 22;
Matches	6.8%	Pred. No. 5.9e-152;
Matches	368;	Conservative 78; Mismatches 96; Indels 26; Gaps
QY	1 MNKLYGNILSENAAPSDDESFINKDAKIPVSGPFVLKTKWAFVDCPDESWALKAIKALSKK	60
Db	40 MNKLYGNLSPAVTADDRLQLGDKPLAQVLLKGSKAYFDYDQPDNWA,TRAITLSSK	99
QY	61 TELHGKPTEVEHSPVKORIRKLQIRNTPHLQEVLDLLVQGYGVSCEQWNTDSEA	120
Db	100 VELHGKMEVDYSKSKLRSRKIQNRNTPHLQEVLDLLAQTGVENTQVNNTDETA	159
QY	121 VVNVTIVSKOARQALDKLNGFOLNETLKAYIDPEMAQQNPLQOPQRGRGLQRGSS	180
Db	160 VVNVTIVTREBREAKLAMELGSQHOFENYSFKS1QSPDEEVSSEPPVQ - RAQR - -GDHSR	215
QY	181 RQG-SPGSVSKQKPODLPLRLVPTQFGVLTGKGATRNTKQTKSYDVRKENSA	239
Db	216 EGHARGGTQARQDIFPLRILVPTQFGVLTGKGELTKITKQTKOSRVJTHRENSGA	275
QY	240 AEKSITLSTPGETSAACKSLBIMIKEAQDIKTFEEIDLKLAHNMFVGLIGKEGRNL	299
Db	276 AEKPVTHATPEGMSEAGCRMILEMOKADESTKLAEELPLKLAHNGLVGRIGKEGRNL	335
QY	300 KKEEQDTDKTISPLQETLNPERTITYGVNTCAAAEEMKKREPENELASM	359
Db	336 KKIEHETGKTISLQLDLSINPERITVGTBEACASEEIMKKLREAFENDMLAVN	395
QY	360 LOAHLPLGLNNAQGLFPPTSGLMPPTSGPPSAMPPYQFEQ - SEETVIOFIPALSVG	418
Db	396 TFS - - - - - GYF - - - - - S - SLYTHHQFGPFFHHSYQEOPQINNFIFPTQAVG	435
QY	419 ATIGKQOQHQLKQLSRFGASIKIAPAAFPDAKVRNYITGPPEAQFKAQGRYIGKKEEN	478
Db	436 ATIGKGAHKQLKQLARFGASIKIAPAGGPDSERAVIITGPPEAQFKAQGRFQKKEEN	495
QY	479 FVSPKVEYKLEAHIVPSPAGHVIGKGKPYNELQNLSSAEEVYVPDQTDENOOVVK	538
Db	496 FVSPKVEYKLEAHIVPSPAGHVIGKGKPYNELQNLSSAEEVYVPDQTDENEEVIVR	555
QY	539 TGHFHZCQVAQRKIOEITQVKHQOO	566
Db	556 TIGHFFASOQRKIREIVQVKQEOK	583

RESULT 14  
 AAM93826  
 ID AAM93826 standard; Protein; 319 AA.  
 XX  
 AC AAM93826;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human polypeptide, SEQ ID NO: 3887.  
 XX  
 KW Human; full length cDNA; cDNA synthesis; oligo-capping.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP113094-A2.  
 XX  
 PD 05-SEP-2001.  
 XX  
 PF 07-JUL-2000; 2000EP-0114089.  
 XX  
 PR 08-JUL-1999; 99JP-0194486.  
 PR 11-JAN-2000; 2000JP-0116774.  
 PR 02-MAY-2000; 2000JP-0183765.  
 XX  
 PA (HELI-) HELIX RES INST.

PI	Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;	XX	XX	XX	OS	Homo sapiens.
XX	WPI; 2001-52455/58.	DR	N-PSDB; AA94162.	XX	PN	WO200155322-A2.
XX	PR	XX	XX	XX	PD	02-AUG-2001.
PT	830 Primers useful for synthesizing full length cDNA clones and their use in genetic manipulation	XX	XX	XX	XX	XX
PR	XX	XX	XX	XX	XX	2001IWO-0501341.
PS	Claim 8; SEQ ID NO 3897; 1380pp + sequence listing; English.	XX	XX	XX	XX	XX
XX	The invention relates to primers for synthesising full length cDNA clones. 830 cDNA molecules encoding a human protein have been isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have been determined. Primers for synthesising the full length cDNA are useful for clarifying the function of the protein encoded by the cDNA. The full length clones were obtained by construction of full length enriched cDNA libraries that were synthesised by the oligo-capping method. The primers enable the production of the full length cDNA easily without any special methods. The present sequence is a polypeptide encoded by a full length human cDNA of the invention.	CC	CC	CC	CC	CC
CC	Note: The sequence data for this patent did not form part of the printed specification but was obtained in CD-ROM format directly from EPO.	CC	CC	CC	CC	CC
XX	Sequence 319 AA:	41.7%	Score 1232; DB 22; Length 319;	Best Local Similarity 76.5%; Prid. No 1.5e-97;	Gaps 5;	XX
XX	Matches 247; Conservative 32; Mismatches 36; Indels 8;	XX	XX	XX	XX	XX
Qy	260 ILEIMHKEAQDIDKTFTEIPKLIAHNNGVGRIGEGRNLKKIEQDTDTKTTISPLQELT 319	;	;	;	;	;
Db	2 ILEIMHKEAQDIDKTFTEIPKLIAHNNGVGRIGEGRNLKKIEQDTDTKTTISLQLDT 61	;	;	;	;	;
Qy	320 LYNPRTITVKGVETCAKEDEIMKKIRESYEDTASNNLQLHILGFLPPT 379.	;	;	;	;	;
Db	62 LYNPRTITVKGALENCCRAGEIMKKVRAYENDVAANSLSQSHLIGLNLAAVGLFPAS 121	;	;	;	;	;
Qy	380 SGMPPTSGPPSAMT-PPYPOFEQS-ETEEVHOFTIPALSGVATIGKGQHQHIKOLSRAFG 436	;	;	;	;	;
Db	122 SSAVPP--PPSSPTGAPYSSFMQPEQRVQFIPRAVGTLIGKQHQIKQLSRFPAS 178	;	;	;	;	;
Qy	437 ASIKFAPAEIPDAKVRVWVITGPPAEOFQAGRIGYKIKEENVSPKEEVKLEAHIRVPS 496	;	;	;	;	;
Db	179 ASIKFAPPETPDSKVWVITGPPAEOFQAGRIGYKIKEENFFGPKEEVKLEAHIRVPA 238	;	;	;	;	;
Qy	497 FAAGRIVIGKGGKTVNELLONLSSAENVVPRODTPDNDVVKTGTGHYACOYAORKIQEI 556	;	;	;	;	;
Db	239 SAAGRIVIGKGGKTVNELLONLTAEVVTPDNTQVPRDQTPDNDVVKTGTGHYASQMAQRKIRD 298	;	;	;	;	;
Qy	557 LTVQYQHOOKQKALOSGPQPSRK 579	;	;	;	;	;
Db	299 LAQVKQ-QRQKG-QSNWQARRK 319	;	;	;	;	;
RESULT 15	XX	XX	XX	XX	XX	XX
AAU16161	07 - NOV - 2001 (first entry)	XX	XX	XX	XX	XX
ID	Human novel secreted protein, Seq ID 1114.	XX	XX	XX	XX	XX
AAU16161 standard; Protein: 261 AA.	AC	AC	AC	AC	AC	AC
AAU16161:	XX	XX	XX	XX	XX	XX
XX	07 - NOV - 2001 (first entry)	XX	XX	XX	XX	XX
DE	Human novel secreted protein, Seq ID 1114.	XX	XX	XX	XX	XX
XX	07 - NOV - 2001 (first entry)	XX	XX	XX	XX	XX
KW	Human; immunosuppressive; antiarthritic; antirheumatic;	XX	XX	XX	XX	XX
KW	cytostatic; cardiotonic; vasotropin; cerebroprotective; nootropic;	XX	XX	XX	XX	XX
KW	neuroprotective; antibacterial; virucide; fungicide; opthalmological;	XX	XX	XX	XX	XX
KW	vulnary; secreted protein; rheumatoid arthritis;	XX	XX	XX	XX	XX
KW	hyperpoliisfer disorder; cardiovascular disorder; cardiac arrest;	XX	XX	XX	XX	XX
KW	cerebrovascular disorder; cerebral ischaemia; angiogenesis;	XX	XX	XX	XX	XX
KW	nervous system disorder; Alzheimer's disease; infection; ocular disorder;	XX	XX	XX	XX	XX
KW	corneal infection; wound healing; epithelial cell proliferation;	XX	XX	XX	XX	XX
KW	skin ageing; food additive; preservative; antiproliferative.	XX	XX	XX	XX	XX

PR 26-SEP-2000; 20000US-0233484.  
 PR 27-SEP-2000; 20000US-0235834.  
 PR 27-SEP-2000; 20000US-0235836.  
 PR 29-SEP-2000; 20000US-0236327.  
 PR 29-SEP-2000; 20000US-0236367.  
 PR 29-SEP-2000; 20000US-0236368.  
 PR 29-SEP-2000; 20000US-0236369.  
 PR 29-SEP-2000; 20000US-0236370.  
 PR 02-OCT-2000; 20000US-0236802.  
 PR 02-OCT-2000; 20000US-0237037.  
 PR 02-OCT-2000; 20000US-0237038.  
 PR 02-OCT-2000; 20000US-0237039.  
 PR 02-OCT-2000; 20000US-0237040.  
 PR 13-OCT-2000; 20000US-0239935.  
 PR 13-OCT-2000; 20000US-0239937.  
 PR 20-OCT-2000; 20000US-0241809.  
 PR 20-OCT-2000; 20000US-0241860.  
 PR 01-NOV-2000; 20000US-0241821.  
 PR 08-NOV-2000; 20000US-0241785.  
 PR 20-OCT-2000; 20000US-0241786.  
 PR 20-OCT-2000; 20000US-0241787.  
 PR 20-OCT-2000; 20000US-0241808.  
 PR 20-OCT-2000; 20000US-0241809.  
 PR 08-NOV-2000; 20000US-0241826.  
 PR 01-NOV-2000; 20000US-0244617.  
 PR 08-NOV-2000; 20000US-0244618.  
 PR 08-NOV-2000; 20000US-0244674.  
 PR 08-NOV-2000; 20000US-0244675.  
 PR 08-NOV-2000; 20000US-0244676.  
 PR 08-NOV-2000; 20000US-0244677.  
 PR 08-NOV-2000; 20000US-0244678.  
 PR 08-NOV-2000; 20000US-0246523.  
 PR 08-NOV-2000; 20000US-0246524.  
 PR 08-NOV-2000; 20000US-0246525.  
 PR 08-NOV-2000; 20000US-0246526.  
 PR 08-NOV-2000; 20000US-0246527.  
 PR 08-NOV-2000; 20000US-0246528.  
 PR 08-NOV-2000; 20000US-0246532.  
 PR 08-NOV-2000; 20000US-0246609.  
 PR 08-NOV-2000; 20000US-0246610.  
 PR 08-NOV-2000; 20000US-0246611.  
 PR 08-NOV-2000; 20000US-0246613.  
 PR 17-NOV-2000; 20000US-0249207.  
 PR 17-NOV-2000; 20000US-0249208.  
 PR 17-NOV-2000; 20000US-0249209.  
 PR 17-NOV-2000; 20000US-0249210.  
 PR 17-NOV-2000; 20000US-0249211.  
 PR 17-NOV-2000; 20000US-0249212.  
 PR 17-NOV-2000; 20000US-0249213.  
 PR 17-NOV-2000; 20000US-0249214.  
 PR 17-NOV-2000; 20000US-0249215.  
 PR 17-NOV-2000; 20000US-0249216.  
 PR 17-NOV-2000; 20000US-0249217.  
 PR 17-NOV-2000; 20000US-0249218.  
 PR 17-NOV-2000; 20000US-0249244.  
 PR 17-NOV-2000; 20000US-0249245.  
 PR 17-NOV-2000; 20000US-0249264.  
 PR 17-NOV-2000; 20000US-0249265.  
 PR 17-NOV-2000; 20000US-0249297.  
 PR 17-NOV-2000; 20000US-0249299.  
 PR 06-DEC-2000; 20000US-0249300.  
 PR 08-DEC-2000; 20000US-0250160.  
 PR 01-DEC-2000; 20000US-0250391.  
 PR 05-DEC-2000; 20000US-0251030.  
 PR 05-DEC-2000; 20000US-0251988.  
 PR 05-DEC-2000; 20000US-0256719.  
 PR 08-DEC-2000; 20000US-0251479.  
 PR 08-DEC-2000; 20000US-0251856.  
 PR 08-DEC-2000; 20000US-0251868.  
 PR 08-DEC-2000; 20000US-0251869.  
 PR 08-DEC-2000; 20000US-0251980.  
 PR 11-DEC-2000; 20000US-0254497.  
 PR 05-JAN-2001; 20010US-0259678.  
 XX

PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI ROSEN CA, Barash SC, Ruben SM;  
 XX  
 WPI: 2001-488783/53.  
 DR N-PSDB; AA326148.  
 XX  
 PT New nucleic acid molecules encoding 461 human secreted proteins for  
 PT diagnosing preventing, treating or ameliorating medical conditions and  
 PT used as food additives or preservatives -  
 XX  
 Claim 11; SEQ ID No 1114; 980sp; English.  
 PS  
 XX  
 XX  
 The invention relates to isolated nucleic acid molecules and their  
 encoded secreted proteins. The nucleic acids and proteins are used to  
 prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
 rabbits, goats, horses, cats, dogs, chickens or sheep. They  
 are also used in diagnosing a pathological condition or susceptibility  
 to a pathological condition. Antibodies to the proteins can also  
 be used in alleviating symptoms associated with the disorders and in  
 diagnostic immunoassays e.g. radioimmunoassays or enzyme linked  
 immunosorbent assays (ELISA). Disorders which are diagnosed or treated  
 include autoimmune diseases e.g. rheumatoid arthritis, cardiovascular disorders  
 e.g. cardiac arrest, cerebrovascular disorders  
 e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.  
 Alzheimer's disease, infections caused by bacteria, viruses and fungi  
 and ocular disorders e.g. corneal infection, and many other  
 disorders listed in the specification. The polypeptides can also  
 be used to aid wound healing and epithelial cell proliferation, to  
 prevent skin ageing due to sunburn, to maintain organs before  
 transplantation, for supporting cell culture of primary tissues, to  
 regenerate tissues and in chemotaxis. The polypeptides can also be used  
 as a food additive or preservative to increase or decrease storage  
 capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
 minerals, cofactors and other nutritional components. The present  
 sequence represents a novel secreted protein of the invention.

Query Match 32.1%; Score 949; DB 22; Length 261;  
 Best Local Similarity 74.5%; Pred. No. 2.6e-73;  
 Matches 193; Conservative 26; Mismatches 32; Indels 8; Gaps 5;

QY 324 ERITVKGNVETCAKEEIMKKIRESYENDIASMNLQAHLPGLPGLNGLFPPTSGMP 383  
 Db 8 ERITVKGAENCCRAEQEMKKVREAYENDVAAMSLSQSHLGLNLAVGLEPASSAV 67

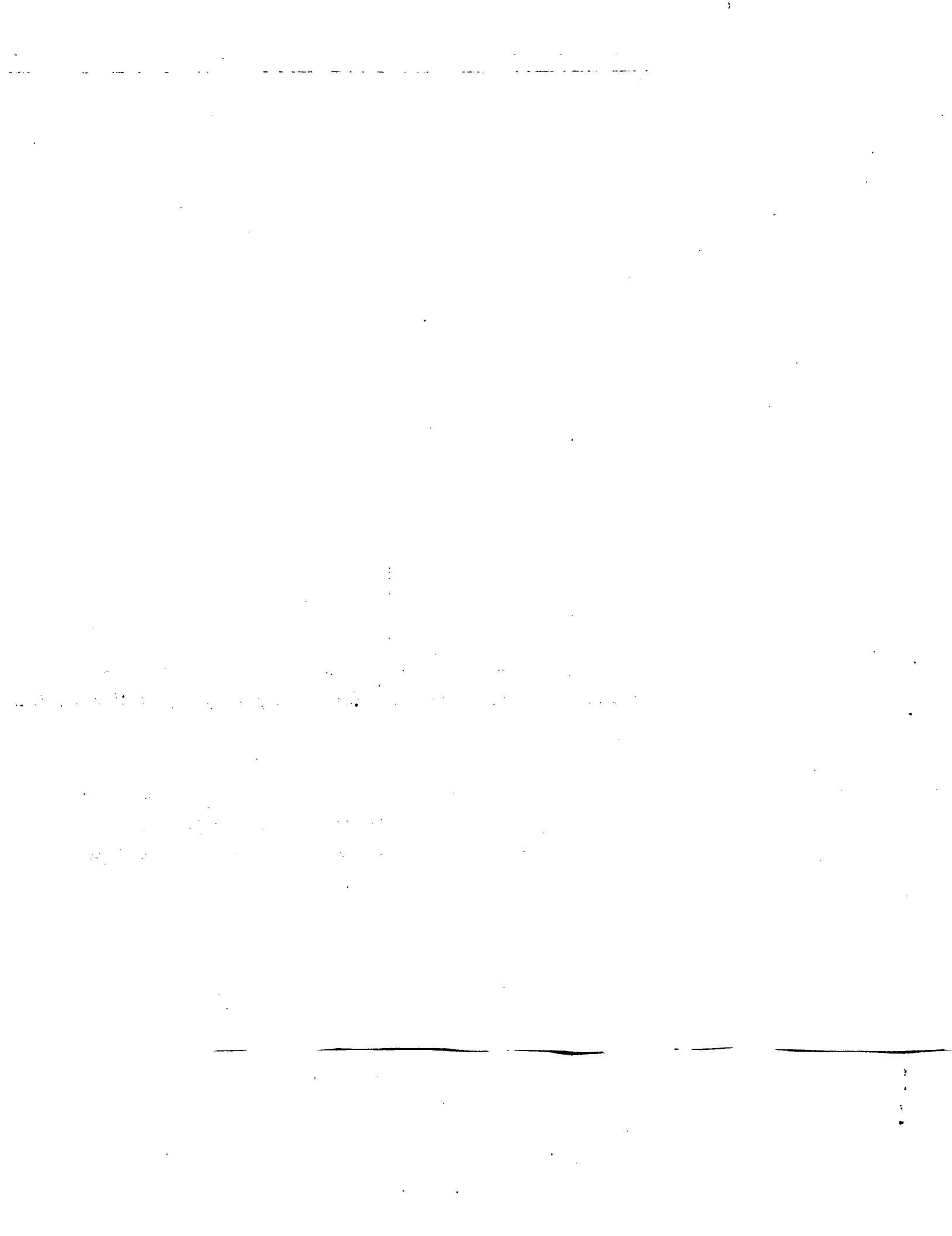
QY 384 PPTSGPPSAMT - PPTPGPEQS - ETETVHQFIPALSVAILKGKQHIKQLSRFGASIK 440  
 Db 68 PP -- PPSSVIGAAFFSSPMQAPQEMVQFIPAQAVAIIGRKQHIKQLSRFASIK 124

QY 441 IAPAPADPKVRYMITGPPEAQPKAQGRIGYKIKEENFSPEKEVILAEHRVPSTAAG 500  
 Db 125 IAPPETPDSKVRYMTITGPPEAQPKAQGRIGYKIKEENFFGPKEEVKLETHIRVPASAG 184

QY 501 RVIGKGKTVNELONLISSABVVPRDQTENDOQVKITGHFYACQVAQRKIQEILTQV 560  
 Db 185 RVIGKGKTVNELONLTAEVVVPRDQIPEENDQVVKIGHFYASQMAQRKIRDIAQV 244

QY 561 KOHOOKQKALOSGPPOSRRK 579  
 Db 245 KQ-OHQKG-QSNOQAQRRK 261

Search completed: April 24, 2003, 16:15:28  
 Job time : 43 secs



GenCore version 5.1.4\_P5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 24, 2003, 16:11:07 : Search time 22 Seconds

(without alignments)

2108.871 Million cell updates/sec

Title: US-p9-897-778-176

Perfect score: 2956

Sequence: 1 MNKLYIGNLSENAAPSDEIES.....VKQHQQQKALQSGPPQSRRK 579

Scoring table: BL21UMG62

Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA.\*

```

1: /cgn2_6/ptodata/1/pubpaas/us08_new_pub.pep:*
2: /cgn2_6/ptodata/1/pubpaas/us07_new_pub.pep:*
3: /cgn2_6/ptodata/1/pubpaas/us06_new_pub.pep:*
4: /cgn2_6/ptodata/1/pubpaas/us07_new_pub.pep:*
5: /cgn2_6/ptodata/1/pubpaas/us07_new_pub.pep:*
6: /cgn2_6/ptodata/1/pubpaas/us07_pubcomb.pep:*
7: /cgn2_6/ptodata/1/pubpaas/us08_pubcomb.pep:*
8: /cgn2_6/ptodata/1/pubpaas/us08_pubcomb.pep:*
9: /cgn2_6/ptodata/1/pubpaas/us09_new_pub.pep:*
10: /cgn2_6/ptodata/1/pubpaas/us09_pubcomb.pep:*
11: /cgn2_6/ptodata/1/pubpaas/us10_new_pub.pep:*
12: /cgn2_6/ptodata/1/pubpaas/us60_pubcomb.pep:*
13: /cgn2_6/ptodata/1/pubpaas/us60_new_pub.pep:*
14: /cgn2_6/ptodata/1/pubpaas/us60_pubcomb.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2956	100.0	579	10	US-09-735-705-176 Sequence 176, App
2	2956	100.0	579	10	US-09-150-116-176 Sequence 176, App
3	2956	100.0	579	10	US-09-897-778-449 Sequence 176, App
4	2943	95.6	579	10	US-09-735-705-348 Sequence 348, App
5	2943	95.6	579	10	US-09-850-716A-348 Sequence 348, App
6	2943	95.6	579	10	US-09-889-778-349 Sequence 346, App
7	2943	95.6	579	10	US-09-897-778-446 Sequence 446, App
8	2943	95.6	579	10	US-09-897-778-449 Sequence 449, App
9	2938	99.4	586	10	US-09-850-716A-427 Sequence 427, App
10	2938	99.4	586	10	US-09-897-778-427 Sequence 427, App
11	2190	74.1	577	10	US-09-873-637-2 Sequence 2, App
12	196.5	65.8	620	10	US-09-764-864-1116 Sequence 1116, App
13	949	32.1	261	10	US-09-764-864-1114 Sequence 1114, App
14	919	31.1	250	10	US-09-764-864-1532 Sequence 1532, App
15	626	21.8	171	10	US-09-764-864-1119 Sequence 1119, App
16	527	19.2	192	10	US-09-764-864-1117 Sequence 1117, App
17	474	16.0	93	10	US-09-64-61-48606 Sequence 48606, App
18	402	13.6	171	10	US-09-764-864-1536 Sequence 1536, App
19	238	8.1	49	10	US-09-873-637-22 Sequence 22, App

## ALIGNMENTS

```

RESULT 1
US-09-735-705-176
; Sequence 176, Application US/09735705
; Patent No. US200300532941
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bagur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LONG CANCER
; FILE REFERENCE: 210121.455C14
; CURRENT APPLICATION NUMBER: US/09/735,705
; CURRENT FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-735-705-176
Query Match 100.% Score 2956; DB 10; Length 579;
; Best Local Similarity 100.%; Pred. No. 1, 7e-217;
; Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MNKLYIGNLSENAAPSDEIESFKDAKIPVSQGPFLVKTGAFVDCPDESNAKALIEALSGK 60
Db 1 MNKLYIGNLSENAAPSDEIESFKDAKIPVSQGPFLVKTGAFVDCPDESNAKALIEALSGK 60
Qy 61 IELHKPKTEVEVSPKRQTRKLQIRNTRPHLOWEVLDSLLVQGVVBSCQVNTDSETA 120
Db 61 IELHKPKTEVEVSPKRQTRKLQIRNTRPHLOWEVLDSLLVQGVVBSCQVNTDSETA 120

```

Query 1.21 VNNTYSSDQAROALDKLNGFOLNETFLKVAIPDIMAQAQNPLQQPRGRGLGORGS 180  
 Database 121 VNNTYSSDQAROALDKLNGFOLNETFLKVAIPDIMAQAQNPLQQPRGRGLGORGS 180.

Query 1.81 RQGSGSVSKQPCDPPLRLVPTQFGAIIGKEGATIRNITKOTQSIDVHKENAGA 240  
 Database 181 RQGSGSVSKQPCDPPLRLVPTQFGAIIGKEGATIRNITKOTQSIDVHKENAGA 240.

Query 2.41 EKSITLSPETGSAAACKSILEMKHEADQIKFTEIPLKLHANNFVGRLIGEGRNLK 300  
 Database 241 EKSITLSPETGSAAACKSILEMKHEADQIKFTEIPLKLHANNFVGRLIGEGRNLK 300.

Query 3.01 KIEQDDTKTISPLQETLYPERTITYKGNEYTCAKAEETIMKIRSYENDIASMNL 360  
 Database 301 KIEQDDTKTISPLQETLYPERTITYKGNEYTCAKAEETIMKIRSYENDIASMNL 360.

Query 3.61 QAHLPGLNINALGLFPSTSMPPTSGPPSAMTPYPOFEOSETETHQFIPALSYGAI 420  
 Database 361 QAHLPGLNINALGLFPSTSMPPTSGPPSAMTPYPOFEOSETETHQFIPALSYGAI 420.

Query 3.61 QAHLPGLNINALGLFPSTSMPPTSGPPSAMTPYPOFEOSETETHQFIPALSYGAI 420  
 Database 361 QAHLPGLNINALGLFPSTSMPPTSGPPSAMTPYPOFEOSETETHQFIPALSYGAI 420.

Query 4.21 IKQGOHIIKOLSRAFASIKIAPAEAPDKVRYAVITGPEAQGRYIGKIKEENFY 480  
 Database 421 IKQGOHIIKOLSRAFASIKIAPAEAPDKVRYAVITGPEAQGRYIGKIKEENFY 480.

Query 4.81 SPKEEVKLEHTRIPSFAGRYIGKOGKTYNELONLSSEAUVVPRQTPENDOVVKIT 540  
 Database 481 SPKEEVKLEHTRIPSFAGRYIGKOGKTYNELONLSSEAUVVPRQTPENDOVVKIT 540.

Query 4.81 SPKEEVKLEHTRIPSFAGRYIGKOGKTYNELONLSSEAUVVPRQTPENDOVVKIT 540  
 Database 481 SPKEEVKLEHTRIPSFAGRYIGKOGKTYNELONLSSEAUVVPRQTPENDOVVKIT 540.

RESULT 3 US-09-897-778-176

Sequence 1.76 Application US/09897778  
 Patent No. US2002014743A1  
 GENERAL INFORMATION:  
 APPLICANT: Wang, Tongtong  
 APPLICANT: Marnerakis, Margarita  
 APPLICANT: Fanger, Gary R.  
 APPLICANT: Vediwick, Thomas S.  
 APPLICANT: Carter, Darren  
 APPLICANT: Watanaabe, Yoshihiro  
 APPLICANT: Henderson, Robert A.  
 APPLICANT: Peckham, David W.  
 APPLICANT: Fanger, Neil  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
 FILE REFERENCE: 210121\_455C16  
 CURRENT APPLICATION NUMBER: US/09/897,778  
 CURRENT FILING DATE: 2001-06-28  
 NUMBER OF SEQ ID NOS: 467  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO 176  
 LENGTH: 579  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-897-778-176

Query Match 1.00 %; Score 2956; DB 10; Length 579;  
 Best Local Similarity 100.0%; Pred. No. 1\_7e-217; Mismatches 0; Indels 0; Gaps 0;  
 Matches 579; Conservative 0; Mismatches 0;

Query 1 MNKLIGNLSENAAPSLESIFDAKIPSGPLVKRYAFVDCPDESWALKALEALSGK 60  
 Database 1 MNKLIGNLSENAAPSLESIFDAKIPSGPLVKRYAFVDCPDESWALKALEALSGK 60

Query 61 IELHGKPIEHSVPKRORIKLQINRNPFLQEVLDLSUVOGVVSECEQVNTDSETA 120  
 Database 61 IELHGKPIEHSVPKRORIKLQINRNPFLQEVLDLSUVOGVVSECEQVNTDSETA 120

Query 61 IELHGKPIEHSVPKRORIKLQINRNPFLQEVLDLSUVOGVVSECEQVNTDSETA 120  
 Database 61 IELHGKPIEHSVPKRORIKLQINRNPFLQEVLDLSUVOGVVSECEQVNTDSETA 120

Query 121 VNNTYSSDQAROALDKLNGFOLNETFLKVAIPDIMAQNPLQQPRGRGLGORGS 180  
 Database 121 VNNTYSSDQAROALDKLNGFOLNETFLKVAIPDIMAQNPLQQPRGRGLGORGS 180.

Query 181 RQGSPGVSKQPCDPPLRLVPTQFGAIIGKEGATIRNITKOTQSIDVHKENAGA 240  
 Database 181 RQGSPGVSKQPCDPPLRLVPTQFGAIIGKEGATIRNITKOTQSIDVHKENAGA 240.

Query 181 RQGSPGVSKQPCDPPLRLVPTQFGAIIGKEGATIRNITKOTQSIDVHKENAGA 240  
 Database 181 RQGSPGVSKQPCDPPLRLVPTQFGAIIGKEGATIRNITKOTQSIDVHKENAGA 240.

QY 241 EKSTITLSTPEGTSAACKSILEIMHKBAODIKFTEEIPILKILAHNNFYGLIGREGRNLK 300  
 DB 241 EKSTITLSTPEGTSAACKSILEIMHKBAODIKFTEEIPILKILAHNNFYGLIGREGRNLK 300  
 QY 301 KIEQDTDKITISPLQELTYNPERTVKGNEVETAKAEEEIMKKIRSYENDIASMNL 360  
 DB 301 KIEQDTDKITISPLQELTYNPERTVKGNEVETAKAEEEIMKKIRSYENDIASMNL 360  
 QY 361 QAHLIPGLNINALGLFPPTSGMPPTSGPPSAMTPYPPQESESETETWQFIPALSYGA 420  
 DB 361 QAHLIPGLNINALGLFPPTSGMPPTSGPPSAMTPYPPQESESETETWQFIPALSYGA 420  
 QY 421 IGKOGHIIKOLSRFGASISKIAAEADAKTRMVITITGPPEAQFAQGRIGYGIKEENFV 480  
 DB 421 IGKOGHIIKOLSRFGASISKIAAEADAKTRMVITITGPPEAQFAQGRIGYGIKEENFV 480  
 QY 481 SPKEEVKLEAHIRVPSAAGRIVKGKTVNELQNLSSAEVVPRDQTIDENDQVVVKIT 540  
 DB 481 SPKEEVKLEAHIRVPSAAGRIVKGKTVNELQNLSSAEVVPRDQTIDENDQVVVKIT 540  
 QY 541 GHFYACQVAQPKIOEILTQVKHQOORALOSCPPOSRK 579  
 DB 541 GHFYACQVAQPKIOEILTQVKHQOORALQSGPPQRK 579  
 QY 541 GHFYACQVAQPKIOEILTQVKHQOORAKALQSGPPQRK 579  
 DB 541 GHFYACQVAQPKIOEILTQVKHQOORAKALQSGPPQRK 579

RESULT 4  
 ; Sequence 348, Application US/09735705  
 ; PATENT NO. US2002005229A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Tongtong  
 ; APPLICANT: Fan, Liqun  
 ; APPLICANT: Kalos, Michael D.  
 ; APPLICANT: Bangir, Chaitanya S.  
 ; APPLICANT: Hosken, Nancy  
 ; APPLICANT: Fanger, Gary R.  
 ; APPLICANT: Li, Samuel X.  
 ; APPLICANT: Wang, Aljun  
 ; APPLICANT: Skelley, Yasir A.W.  
 ; APPLICANT: Henderson, Robert A.  
 ; APPLICANT: McNeill, Patricia D.  
 ; TIME OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
 ; FILE REFERENCE: 210121\_455C14  
 ; CURRENT APPLICATION NUMBER: US/09/735,705  
 ; CURRENT FILING DATE: 2000-12-12  
 ; SOFTWARE: FastSEQ for Windows Version 3.0  
 ; SEQ ID NO: 348  
 ; LENGTH: 579  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-735-705-348

Query Match 99.68; Score 2943; DB 10; Length 579;  
 Best Local Similarity 99.7%; Pred. No. 1.7e-216;  
 Matches 577; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MNKLIGNLSNAAFSDLESIFDKATIVPSGPFPLVKTGAFVDCPDESVALKATEALSGK 60  
 DB 1 MNKLIGNLSNAAFSDLESIFDKATIVPSGPFPLVKTGAFVDCPDESVALKATEALSGK 60  
 QY 61 IELHGKPTEVEHSVPKRQRIKLOJRNPHQWEYDLSLIVQGVVSESEQNTDSETA 120  
 DB 61 IELHGKPTEVEHSVPKRQRIKLOJRNPHQWEYDLSLIVQGVVSESEQNTDSETA 120  
 QY 121 VNVYSSKDOARQALQDNGFOLNFNTKVALYIPDEMAQNPQLQPRGRGLGORGSS 180  
 DB 121 VNVYSSKDOARQALQDNGFOLNFNTKVALYIPDETAQNPQLQPRGRGLGORGSS 180  
 QY 181 ROGSPGSYSKOPCDJPLRLVYPTOFVGATIGKFGATLNTKOTOSIDYHRKENAGAA 240  
 DB 181 ROGSPGSYSKOPCDJPLRLVYPTOFVGATIGKFGATLNTKOTOSIDYHRKENAGAA 240  
 QY 241 EKSTITLSTPEGTSAACKSILEIMHKBAODIKFTEEIPILKILAHNNFYGLIGREGRNLK 300  
 DB 241 EKSTITLSTPEGTSAACKSILEIMHKBAODIKFTEEIPILKILAHNNFYGLIGREGRNLK 300  
 QY 301 KIEQDTDKITISPLQELTYNPERTVKGNEVETAKAEEEIMKKIRSYENDIASMNL 360  
 DB 301 KIEQDTDKITISPLQELTYNPERTVKGNEVETAKAEEEIMKKIRSYENDIASMNL 360  
 QY 361 QAHLIPGLNINALGLFPPTSGMPPTSGPPSAMTPYPPQESESETETWQFIPALSYGA 420  
 DB 361 QAHLIPGLNINALGLFPPTSGMPPTSGPPSAMTPYPPQESESETETWQFIPALSYGA 420  
 QY 421 SPKEEVKLEAHIRVPSAAGRIVKGKTVNELQNLSSAEVVPRDQTIDENDQVVVKIT 540  
 DB 421 SPKEEVKLEAHIRVPSAAGRIVKGKTVNELQNLSSAEVVPRDQTIDENDQVVVKIT 540  
 QY 481 GHFYACQVAQPKIOEILTQVKHQOORALOSCPPOSRK 579  
 DB 481 GHFYACQVAQPKIOEILTQVKHQOORAKALQSGPPQRK 579  
 QY 541 GHFYACQVAQPKIOEILTQVKHQOORAKALQSGPPQRK 579

RESULT 5  
 US-09-850-716A-348  
 ; Sequence 348, Application US/09850716A  
 ; Patent No. US20020115139A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kalos, Michael D.  
 ; APPLICANT: McNeill, Patricia D.  
 ; APPLICANT: Reiter, Marc W.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
 ; FILE REFERENCE: 210121\_455C15  
 ; CURRENT APPLICATION NUMBER: US/09/850,716A  
 ; CURRENT FILING DATE: 2001-05-07  
 ; NUMBER OF SEQ ID NO: 440  
 ; SOFTWARE: FastSEQ for Windows Version 3.0  
 ; SEQ ID NO: 348  
 ; LENGTH: 579  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-850-716A-348

Query Match 99.68; Score 2943; DB 10; Length 579;  
 Best Local Similarity 99.7%; Pred. No. 1.7e-216;  
 Matches 577; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MNKLIGNLSNAAFSDLESIFDKATIVPSGPFPLVKTGAFVDCPDESVALKATEALSGK 60  
 DB 1 MNKLIGNLSNAAFSDLESIFDKATIVPSGPFPLVKTGAFVDCPDESVALKATEALSGK 60  
 QY 61 IELHGKPTEVEHSVPKRQRIKLOJRNPHQWEYDLSLIVQGVVSESEQNTDSETA 120  
 DB 61 IELHGKPTEVEHSVPKRQRIKLOJRNPHQWEYDLSLIVQGVVSESEQNTDSETA 120  
 QY 121 VNVYSSKDOARQALQDNGFOLNFNTKVALYIPDEMAQNPQLQPRGRGLGORGSS 180  
 DB 121 VNVYSSKDOARQALQDNGFOLNFNTKVALYIPDETAQNPQLQPRGRGLGORGSS 180  
 QY 181 ROGSPGSYSKOPCDJPLRLVYPTOFVGATIGKFGATLNTKOTOSIDYHRKENAGAA 240  
 DB 181 ROGSPGSYSKOPCDJPLRLVYPTOFVGATIGKFGATLNTKOTOSIDYHRKENAGAA 240  
 QY 241 EKSTITLSTPEGTSAACKSILEIMHKBAODIKFTEEIPILKILAHNNFYGLIGREGRNLK 300  
 DB 241 EKSTITLSTPEGTSAACKSILEIMHKBAODIKFTEEIPILKILAHNNFYGLIGREGRNLK 300  
 QY 301 KIEQDTDKITISPLQELTYNPERTVKGNEVETAKAEEEIMKKIRSYENDIASMNL 360

RESULT 6

US-09-897-778-348

; Sequence 348, Application US/09897778

; Patent No. US2002014714A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Marnerakis, Margarita

; APPLICANT: Fauder, Gary R.

; APPLICANT: Vedick, Thomas S.

; APPLICANT: Carter, Darick

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Henderson, Robert A.

; APPLICANT: Peckham, David W.

; APPLICANT: Ranger, Neil

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121\_455C16

; CURRENT APPLICATION NUMBER: US/09-897-778

; CURRENT FILING DATE: 2001-06-28

; NUMBER OF SEQ ID NOS: 467

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO: 348

; LENGTH: 579

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-897-778-348

Query Match 99.6%; Score 2943; DB 10; Length 579;

Best Local Similarity 99.7%; Prod. No. 1.7e-216;

Matches 577; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MNKLYIGNI&ENAAPSIDESTIFDKAKIPSPFLVYKTYGAFYDPCPDESWALKATEALSGK 60

Db 1 MNKLYIGNI&ENAAPSIDESTIFDKAKIPSPFLVYKTYGAFYDPCPDESWALKATEALSGK 60

Qy 61 TELHGPPIEVSHPKRQKRQLQINNIPPHLOWEVLSSLYQGVYSECVNTDSETA 120

Db 61 TELHGPPIEVSHPKRQKRQLQINNIPPHLOWEVLSSLYQGVYSECVNTDSETA 120

Qy 121 VVNNTSSKQDQARQDALDINGFOLENTLKVAYIDEMAAQQNPLQQRGRGLGGRGS 180

Db 121 VVNNTSSKQDQARQDALDINGFOLENTLKVAYIDETAAQQNPLQQRGRGLGGRGS 180

Qy 181 ROGSPGSVKQPKCDPLRLLYPTQFGAIIKGEGATIRNITKOTSKIDYHRKENAGAA 240

Db 181 ROGSPGSVKQPKCDPLRLLYPTQFGAIIKGEGATIRNITKOTSKIDYHRKENAGAA 240

Qy 241 EKSITLSTPEGTSAACKSILEMHRKAQDIKTFTEIPLKLTLAHNNFGLIGKEGRNLK 300

Db 241 EKSITLSTPEGTSAACKSILEMHRKAQDIKTFTEIPLKLTLAHNNFGLIGKEGRNLK 300

Qy 301 KIEQDTDTKITISPLQETLYNPERTIVKGNVERCAKAEEIMKKIRESYENDIASMNL 360

Db 301 KIEQDTDTKITISPLQETLYNPERTIVKGNVERCAKAEEIMKKIRESYENDIASMNL 360

Qy 361 QAHLIGLNNAALGLFPPTSCHMPPPSGPPSAMPPYPOFEQSETETHQFTPALSGAI 420

Db 361 QAHLIGLNNAALGLFPPTSCHMPPPSGPPSAMPPYPOFEQSETETHQFTPALSGAI 420

Qy 421 IGKQGHIIKOLSRPAGASKIAPAEADAKYRMVITGPPEAQFKAGTRIGKIKEENFV 480

Db 421 IGKQGHIIKOLSRPAGASKIAPAEADAKYRMVITGPPEAQFKAGTRIGKIKEENFV 480

Qy 481 SPKEEKVLEAHTRVPSFAAGRIGKGGTKTNELQNLSSAEVYPROTPDENDQVVKIT 540

Db 481 SPKEEKVLEAHTRVPSFAAGRIGKGGTKTNELQNLSSAEVYPROTPDENDQVVKIT 540

Qy 541 GHFYACQVAQRKIQEILTOXQHOQOKALQSGPQSRRK 579

Db 541 GHFYACQVAQRKIQEILTOXQHOQOKALQSGPQSRRK 579

RESULT 7

US-09-897-778-446

; Sequence 446, Application US/09897778

; Patent No. US2002014714A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Marnerakis, Margarita

; APPLICANT: Fauder, Gary R.

; APPLICANT: Vedick, Thomas S.

; APPLICANT: Carter, Darick

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Henderson, Robert A.

; APPLICANT: Peckham, David W.

; APPLICANT: Ranger, Neil

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121\_455C16

; CURRENT APPLICATION NUMBER: US/09-897-778

; CURRENT FILING DATE: 2001-06-28

; NUMBER OF SEQ ID NOS: 467

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO: 446

; LENGTH: 579

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-897-778-446

Query Match 99.6%; Score 2943; DB 10; Length 579;

Best Local Similarity 99.7%; Pred. No. 1.7e-216;

Matches 577; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MNKLYIGNI&ENAAPSIDESTIFDKAKIPSPFLVYKTYGAFYDPCPDESWALKATEALSGK 60

Db 1 MNKLYIGNI&ENAAPSIDESTIFDKAKIPSPFLVYKTYGAFYDPCPDESWALKATEALSGK 60

Qy 61 TELHGPPIEVSHPKRQKRQLQINNIPPHLOWEVLSSLYQGVYSECVNTDSETA 120

Db 61 TELHGPPIEVSHPKRQKRQLQINNIPPHLOWEVLSSLYQGVYSECVNTDSETA 120

Qy 121 VVNNTSSKQDQARQDALDINGFOLENTLKVAYIDEMAAQQNPLQQRGRGLGGRGS 180

Db 121 VVNNTSSKQDQARQDALDINGFOLENTLKVAYIDEMAAQQNPLQQRGRGLGGRGS 180

Qy 181 ROGSPGSVKQPKCDPLRLLYPTQFGAIIKGEGATIRNITKOTSKIDYHRKENAGAA 240

Db 181 ROGSPGSVKQPKCDPLRLLYPTQFGAIIKGEGATIRNITKOTSKIDYHRKENAGAA 240

Qy 241 EKSITLSTPEGTSAACKSILEMHRKAQDIKTFTEIPLKLTLAHNNFGLIGKEGRNLK 300

Db 241 EKSITLSTPEGTSAACKSILEMHRKAQDIKTFTEIPLKLTLAHNNFGLIGKEGRNLK 300

Qy 301 KIEQDTDTKITISPLQETLYNPERTIVKGNVERCAKAEEIMKKIRESYENDIASMNL 360

Db 301 KIEQDTDTKITISPLQETLYNPERTIVKGNVERCAKAEEIMKKIRESYENDIASMNL 360

Oy	361 OAHUJPGLNALNGLFPPTSGMPPPTSGPPSAMTPYPOFEGSETEVHOFIPALSVGAI 4.20	Qy	361 OAHUJPGLNALNGLFPPTSGMPPPTSGPPSAMTPYPOFEGSETEVHOFIPALSVGAI 4.20
Db	361 OAHUJPGLNALNGLFPPTSGMPPPTSGPPSAMTPYPOFEGSETEVHOFIPALSVGAI 4.20	Db	361 OAHUJPGLNALNGLFPPTSGMPPPTSGPPSAMTPYPOFEGSETEVHOFIPALSVGAI 4.20
Oy	421 IKGQOHHQIKLQSRAGASIKIAPAPAIDAKYRMVITTPPEAQFAQRGKIREENV 4.80	Qy	421 IKGQOHHQIKLQSRAGASIKIAPAPAIDAKYRMVITTPPEAQFAQRGKIREENV 4.80
Db	421 IKGQOHHQIKLQSRAGASIKIAPAPAIDAKYRMVITTPPEAQFAQRGKIREENV 4.80	Db	421 IKGQOHHQIKLQSRAGASIKIAPAPAIDAKYRMVITTPPEAQFAQRGKIREENV 4.80
Oy	481 SPKEEYKLEAHIRVSPAGRVIGSGKTVNEQNLSAEVYYVPRDQPENDQTVVKIT 5.40	Qy	481 SPKEEYKLEAHIRVSPAGRVIGSGKTVNEQNLSAEVYYVPRDQPENDQTVVKIT 5.40
Db	481 SPKEEYKLEAHIRVSPAGRVIGSGKTVNEQNLSAEVYYVPRDQPENDQTVVKIT 5.40	Db	481 SPKEEYKLEAHIRVSPAGRVIGSGKTVNEQNLSAEVYYVPRDQPENDQTVVKIT 5.40
Oy	541 GHFYACQAQRKIQEILTQVKOHOQOKALQSEPPSRK 579	Qy	541 GHFYACQAQRKIQEILTQVKOHOQOKALQSEPPSRK 579
Db	541 GHFYACQAQRKIQEILTQVKOHOQOKALQSEPPSRK 579	Db	541 GHFYACQAQRKIQEILTQVKOHOQOKALQSEPPSRK 579
<hr/>			
RESULT 8			
US-09-897-778-449			
Sequence 4.49 Application US/09869778			
GENERAL INFORMATION:			
APPLICANT: Wang, Tongtong			
APPLICANT: Marigarta			
APPLICANT: Fanger, Gary R.			
APPLICANT: Vedick, Thomas S.			
APPLICANT: Carter, Parrick			
APPLICANT: Watanabe, Yoshihiro			
APPLICANT: Henderson, Robert A.			
APPLICANT: Peckham, David W.			
APPLICANT: Fanger, Neil			
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY			
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER			
FILE REFERENCE: 210121.455C15			
CURRENT APPLICATION NUMBER: US/09/850,716A			
CURRENT FILING DATE: 2001-05-07			
NUMBER OF SEQ ID NOS: 467			
SOFTWARE: FastSEQ for Windows version 4.0			
SEQ ID NO: 449			
LENGTH: 579			
TYPE: PRX			
ORGANISM: Homo sapiens			
US-09-897-778-449			
Query Match 99.68; Score 2943; DB 10; Length 579;			
Best Local Similarity 99.78; Pred: No. 1.7e-216; Indels 0; Gaps 0;			
Matches 577; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Query Match 99.68; Score 2943; DB 10; Length 579;			
Best Local Similarity 99.78; Pred: No. 1.7e-216; Indels 0; Gaps 0;			
Matches 577; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Oy	1 MNKLYGNLSENAPSDLESIFDKAIDPKVSPGPFLVKTGYAFVDCPDESWALKATEALSK 60	Qy	1 MNKLYGNLSENAPSDLESIFDKAIDPKVSPGPFLVKTGYAFVDCPDESWALKATEALSK 61
Db	1 MNKLYGNLSENAPSDLESIFDKAIDPKVSPGPFLVKTGYAFVDCPDESWALKATEALSK 60	Db	9 MNKLYGNLSENAPSDLESIFDKAIDPKVSPGPFLVKTGYAFVDCPDESWALKATEALSK 61
Oy	61 IELHGKPTEYHSVPKRORIKLQNRNTPHLQNEVLDSLLIVQYCVSCEVNDSETA 120	Qy	62 ELHGPIVEHESVPKRORIKLQNRNTPHLQNEVLDSLLIVQYCVSCEVNDSETA 121
Db	61 IELHGKPTEYHSVPKRORIKLQNRNTPHLQNEVLDSLLIVQYCVSCEVNDSETA 120	Db	69 ELHGPIVEHESVPKRORIKLQNRNTPHLQNEVLDSLLIVQYCVSCEVNDSETA 128
Oy	121 VVNTYSSKDOARQALDKLNGFOLENFTLKVAYTFDENAACONPLQQRGRGLQROSS 180	Qy	122 VVNTYSSKDOARQALDKLNGFOLENFTLKVAYTFDENAACONPLQQRGRGLQROSS 181
Db	121 VVNTYSSKDOARQALDKLNGFOLENFTLKVAYTFDENAACONPLQQRGRGLQROSS 180	Db	129 VVNTYSSKDOARQALDKLNGFOLENFTLKVAYTFDENAACONPLQQRGRGLQROSS 188
Oy	121 VVNTYSSKDOARQALDKLNGFOLENFTLKVAYTFDENAACONPLQQRGRGLQROSS 180	Qy	182 QOSPSSVKQKCDPLRBLVTOVCAIGSCATNUTOTQSQIDVHRKENAGE 241
Db	121 VVNTYSSKDOARQALDKLNGFOLENFTLKVAYTFDENAACONPLQQRGRGLQROSS 180	Db	189 QOSPSSVKQCDPLRBLVTOFGAIGBTAITNTOTQSQIDVHRKENAGE 248
Oy	181 RGSPGSVSKOPKCDPLRBLVTOFGAIGKESATIRTKQPSIDVHRKENASA 240	Qy	242 KSITLSPPEGTSAACKSITLEIMHEAOIKEETPLKLAHNHFVGRLIKEGENLKK 301
Db	181 RGSPGSVSKOPKCDPLRBLVTOFGAIGKESATIRTKQPSIDVHRKENASA 240	Db	249 KSITLSPPEGTSAACKSITLEIMHEAOIKEETPLKLAHNHFVGRLIKEGENLKK 308
Oy	241 EKSITLSTPCTGAACKSITLEIMHEAOIKEETPLKLAHNHFVGRLIKEGRNLK 300	Qy	302 LEQDDTKITSPLQBLTLYNPERITYKGNEYETAKEEEMKIRESYENDIASMNQ 361
Db	241 EKSITLSTPCTGAACKSITLEIMHEAOIKEETPLKLAHNHFVGRLIKEGRNLK 300	Db	309 LEQDDTKITSPLQBLTLYNPERITYKGNEYETAKEEEMKIRESYENDIASMNQ 368
Oy	301 KTEDDTDKITISPLQBLTLYNPERITYKGNEYETAKEEEMKIRESYENDIASMNQ 360	Qy	362 AHLPGLNALNGLFPPTSGMPPSGPSAMTPYPOFEGSETEVHOFIPALSVGAI 4.21
Db	301 KTEDDTDKITISPLQBLTLYNPERITYKGNEYETAKEEEMKIRESYENDIASMNQ 360	Db	369 AHLPGLNALNGLFPPTSGMPPSGPSAMTPYPOFEGSETEVHOFIPALSVGAI 4.28
Oy	301 KIEDDTKITSPLQBLTLYNPERITYKGNEYETAKEEEMKIRESYENDIASMNQ 360	Qy	422 GRQGHQIKLQSRAGASIKIAPAEADAKVRKVIITGPPEAQFAQRGKIREBNFVS. 481
Db	301 KIEDDTKITSPLQBLTLYNPERITYKGNEYETAKEEEMKIRESYENDIASMNQ 360	Db	429 GRQGHQIKLQSRAGASIKIAPAEADAKVRKVIITGPPEAQFAQRGKIREBNFVS. 488

RESULT 10

Qy 482 PKEVKLEAHIRVPSAAGRVIGKGGKTVNELQNLSSAEYVPRDOPDENDQVVKITG 541  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Db 489 PKEVKLEAHIRVPSAAGRVIGKGGKTVNELQNLSSAEYVPRDOPDENDQVVKITG 548  
 Qy 542 HRYACOVARKTOEILTYQKORQALQSPPSRK 579  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Db 549 HRYACOVARKTOEILTYQKORQALQSPPSRK 586  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I

RESULT 11

US-09-878-637-2

Sequence 2, Application US/09873637  
 ; Patent No. US2000061543A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ROSS, Jeffrey  
 ; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN  
 ; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN  
 ; FILE REFERENCE: 960296\_95131  
 ; CURRENT APPLICATION NUMBER: US/09/873, 637  
 ; CURRENT FILING DATE: 2001-06-04  
 ; NUMBER OF SEQ ID NOS: 46  
 ; SOFTWARE: Patentin Ver. 2.0  
 ; SEQ ID NO: 2  
 ; LENGTH: 577  
 ; TYPE: PR  
 ; ORGANISM: Mus musculus  
 US-09-878-637-2

Query Match Similarity 74.1%; Score 2190; DB 10; Length 577;  
 Best Local Similarity 74.1%; Pred. No. 5.1e-18; Mismatches 74; Indels 16; Gaps 8;

Matches 434; Conservative 62; Mismatches 74; Indels 16; Gaps 8;

Qy 1 MKLYIGNLSENAAPSDLESIFDKAPIVSGPFLVKTGYAFVDCPDPESWALKATEALSGK 60  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Db 1 MNKLIGNLNESTPADEKVFREHKTSISGPFLVKTGYAFVDCPDPESWALKATEALSGK 60  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Qy 61 IELHAKPIVEEVSPKQRTRKIQNIPPHOMELDSDLLVQYGVESCDWVNTDSETA 120  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Db 61 VEQGRMLMEHSVPKQRKIQNIPPHOMELDSDLLVQYGVESCDWVNTDSETA 120  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Qy 121 VNVVYSSDQARQALDQKNGFOLENTFLKVATIPDMAAQNPLOPQRGR-SLGORGs 179  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Db 121 VNVVYSSDQARQALDQKNGFOLENTFLKVATIPDMAAQNPLOPQRGR-SLGORGs 179  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Qy 180 SROGSP--GSYSPCOPDPLRLLPTQFGVAGLIGKEGATINRNTKOTOSKIDVHRKEN 236  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Db 180 SROGSP--GSYSPCOPDPLRLLPTQFGVAGLIGKEGATINRNTKOTOSKIDVHRKEN 236  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Qy 177 PROGSEVAGAPAKOPVDPDPLRLLPTQFGVAGLIGKEGATINRNTKOTOSKIDVHRKEN 236  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Db 177 PROGSEVAGAPAKOPVDPDPLRLLPTQFGVAGLIGKEGATINRNTKOTOSKIDVHRKEN 236  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Qy 237 AGAAEKSTIUSPBTSAKSILEIMKEDADIREKTEIPKILAHNNFVGVLIGKEG 296  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Db 237 AGAAEKSTIUSPBTSAKSILEIMKEDADIREKTEIPKILAHNNFVGVLIGKEG 296  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Qy 297 RULKIEQDTKITSPIELETLYNPERTITVKGNEYETCAKEEEIMKRTSYENDIA 356  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Db 297 RULKIEQDTKITSPIELETLYNPERTITVKGNEYETCAKEEEIMKRTSYENDIA 356  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Qy 357 'SNLQAHLPGLNANLGLFPPTSGMPPTSGPSAMT - PPYPOFOQS - ETENVHOFTIP 413  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Db 357 ANLSQAHLPGLNANLGLFPPTSGMPPTSGPSAMT - PPYPOFOQS - ETENVHOFTIP 413  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Qy 414 ALSVGAIIKQGOHJKQLSFASASTIKAPAEPDAKVRAVITGPPEAQPKAGRYGK 473  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Db 414 AGAVAIKQGOHJKQLSFASASTIKAPAEPDAKVRAVITGPPEAQPKAGRYGK 473  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Qy 474 LKEENFVSKKEVKLEAHIRVPSAAGRVIGKGGKTVNELQNLSSAEYVPRDOPDEND 533  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Db 474 LKEENFVSKKEVKLEAHIRVPSAAGRVIGKGGKTVNELQNLSSAEYVPRDOPDEND 533  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Qy 534 QVVKITGHFYACQAQRIQELTQVOKHQOKALQSPPSRK 579  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Db 534 QVVKITGHFYACQAQRIQELTQVOKHQOKALQSPPSRK 579  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Qy 422 GRQGHKIQSFRAGSIIATAEADAVRMTTGGPPAQFRACRIGKIKEENFS 481  
 I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I|||||I  
 Db 429 GRQGHKIQSFRAGSIIATAEADAVRMTTGGPPAQFRACRIGKIKEENFS 488

RESULT 12 US-09-764-864-1116  
 Sequence 1116; Application US/09764864  
 Patent No. US20020132753A1  
 GENERAL INFORMATION:  
 APPLICANT: Rosen et al.  
 TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 FILE REFERENCE: PT223  
 CURRENT APPLICATION NUMBER: US/09/764-864  
 CURRENT FILING DATE: 2001-01-17  
 Prior application data removed - consult PALM or file wrapper  
 NUMBER OF SEQ ID NOS: 1792  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO: 1116  
 LENGTH: 620  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE: SITE  
 NAME/KEY: SITE  
 LOCATION: (533)  
 OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids

US-09-764-864-1116

Query Match 65.8%; Score 1946.5; DB 10; Length 620;  
 Best Local Similarity 65.1%; Pred. No. 2.2e-140;  
 Matches 385; Conservative 81; Mismatches 96; Indels 29; Gaps 8;

Qy 1 MNKLVGNLSENAAPSDETLSEIFDKAPIYSGPFVYKTGTAFYDOPCDESWAALKATEALSGK 60  
 Db 23 MNKLVGNLSPAVTADLRQFLDEKLPLAGQVLKSGTAFYDOPCDESWAALKATEALSGK 82

Qy 61 IELHGKPIEVHSPVKRORTRKLQIRNIPHLQKBEYLDLSLVQYGVVESCEQVNNTSETA 120  
 Db 83 VELHGKPIEVHSPVKRORTRKLQIRNIPHLQKBEYLDLSLVQYGVVESCEQVNNTSETA 142

Qy 121 VVNVTYSSKKOARQALDKLNGQFOLENFTLKAVIYDPEMAAQNPLQQPRGRRGLGCRESS 180  
 Db 143 VVNVTYATREAKTAMEKLQSLGHOFENYSPKSYLIPDEEVSSPSPQ-RAGR-GDHESR 198

Qy 181 RQG-SPGSVSKPDPLPLLYPQFQYEAIIIGRATRNTIKOTSIDYVERKENGA 239  
 Db 199 EGHARGGTGTSQARQDPLFLYLPQFQVIGKREGLTNKNTQTSRVDIRKENGKA 258

Qy 240 AEKSITILSTPEGTSAAKSITLEIMKHEADIKPTEEIPIKILAHNNFYGRIGKEGRNL 299  
 Db 259 AEKPTIHAPEGTSAAKSITLEIMKHEADIKPTEEIPIKILAHNNFYGRIGKEGRNL 318

Qy 300 KKIEQDTDTKITSPLQELTLYNPERTIVKGNYETCAKEEEIMKKIRESYENDIASMN 359  
 Db 319 KKIEHETGKTITISSQDLSTYNPERTIVKGTYEACASEIELMKLREAFENDMLAVN 378

Qy 360 LOAHLIPGLNIALGF-----PPTS-----SGPPPT-----P 396  
 Db 379 QANLIPGLNISALGIFTSGLSPPAGRGAPPAPYHPFTHSGFSLYPHHQFGP 438

Qy 397 YPQFEQ-SETETVHQTPLSQTPEAKTAAEPIKLAETPIKILAHNLVRLIGKEGRNL 455  
 Db 439 FPHHSYYPEDEIVNLIPTDAVGAIIGKRAHKAHILQKLAETPIKILAHNLVRLIGKEGRNL 498

Qy 456 ITGPPAEAQFKAQGRIGKIKEENFYSPKVEKLEAHIRYPSAAGRIVKGKGTVNELQN 515  
 Db 499 ITGPPAEAQFKAQGRIGKIKEENFNPKEVKLEXHIVRESSTAGRVIGGGKTVNELQN 558

Qy 516 LSSAEVVPDQPTDENDQVVKTTGHFYACQVAQRKQHQQ 566  
 Db 559 LTSAEVIVPDRQTPDNEEVIVRIGHFFPSQTAQRKIREIVQQVKQEQOK 609

RESULT 13 US-09-764-864-1114  
 Sequence 1114; Application US/09764864  
 Patent No. US20020132753A1  
 GENERAL INFORMATION:  
 APPLICANT: Rosen et al.  
 TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 FILE REFERENCE: PT223  
 CURRENT APPLICATION NUMBER: US/09/764-864  
 CURRENT FILING DATE: 2001-01-17  
 Prior application data removed - consult PALM or file wrapper  
 NUMBER OF SEQ ID NOS: 1792  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO: 1114  
 LENGTH: 261  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-764-864-1114

Query Match 32.1%; Score 949; DB 10; Length 261;  
 Best Local Similarity 74.5%; Pred. No. 8.5e-65;  
 Matches 32; Mismatches 32; Indels 8; Gaps 5;

Qy 324 ERTITYVGNVENCATAKABEEIMKKIRESYENDIASMLQAHJLPGINLNALGLFPPTSGMP 383  
 Db 68 PP---PPSVTGAAPYSSFMQAPEQEMVQVTFPAQAVGALIGKGQHJKQLSRFASASIK 124

Qy 441 TAPAEAPDAKVRMVLITGPPEAQFKAGRIGKIKEBINFVSPKEEVKLEAHIRVPSFAAG 500  
 Db 125 TAPPETDSKVMVITGPPEAQFKAGRIGKIKEBINFVSPKEEVKLEAHIRVPSFAAG 184

Qy 501 RVIGKGKTVNELQNLSSAEVVWPKTGHFYACQVAQRKIQEILITQV 560  
 Db 185 RVIGKGKTVNELQNLAAEVVWPKTGHFYASQMAQRKIKRDILAQV 244

Qy 561 KQHQQQKALOSGPPOSSRK 579  
 Db 245 KQ-QHQKG-QSNOQAQRK 261

RESULT 14 US-09-74-864-1532  
 Sequence 1532; Application US/09764864  
 Patent No. US20020132753A1  
 GENERAL INFORMATION:  
 APPLICANT: Rosen et al.  
 TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 FILE REFERENCE: PT223  
 CURRENT APPLICATION NUMBER: US/09/764-864  
 CURRENT FILING DATE: 2001-01-17  
 Prior application data removed - consult PALM or file wrapper  
 NUMBER OF SEQ ID NOS: 1792  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO: 1532  
 LENGTH: 250  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE: SITE  
 NAME/KEY: SITE  
 LOCATION: (127)  
 OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids

Qy 31.1%; Score 919; DB 10; Length 250;  
 Best Local Similarity 73.3%; Pred. No. 1.6e-62;  
 Matches 187; Conservative 26; Mismatches 34; Indels 8; Gaps 5;

Qy 328 TVKGNAEVTCATAKABEEIMKKIRESYENDIASMLQAHJLPGINLNALGLFPPTSGMP 387  
 Db 1 TVKGNAEVTCATAKABEEIMKKIRESYENDIASMLQAHJLPGINLNALGLFPPTSGMP 387

QY 388 GPPSAMT-~~T~~PYPQFEQS-ETETVHQFTIPALSYGAIIKGKQGHIKQLSRAGASIKIAP A 444  
Db 59 -PPSVTGAPYSSMOPAQEIQVQPTIPQAVGAIIGKQGHIKQLSRASASITIAPP 117

QY 445 EAPDAKVRVVIITGPPEAQFKAQGRIVYKIKEENFSPEKEVYKLEAHIRVPSFAAGRIVG 504  
Db 118 ETPDSKVRYXIXITGPXPXAFQFKAQGRIVYKIKEENFSPEKEVYKLEAHIRVPSAAGRIVG 177

QY 505 KGGKTVNEDDNLSSAEVYVPRDQTPDENQVYVKTGHEYACQVAQRKIOEILTOVKQHQ 564.  
Db 178 KGGKTVNEDDNLTAAEVVPRDQTPDENQVYVKTGHEYASQMAQRKIRDILAQVHQ-Q 236

QY 565 QOKALQSGPPOSRK 579

Db 237 HQKG-QSNQ-QARK 250

## RESULT 15

; Sequence 1119, Application US/09764864

; Patent No. US20020132753A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PT223

; CURRENT APPLICATION NUMBER: US/09/764,864

; CURRENT FILING DATE: 2001-01-17

; PRIOR application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOs: 1752

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1119

; LENGTH: 171

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-764-864-1119

Query Match 21.2%; Score 626; DB 10; Length 171;

Best Local Similarity 80.5%; Prod. No. 2.1e-40; Mismatches 13; Indexes 0; Gaps 0;

Matches 120; Conservative 16;

; GATIGKQGHIKQLSRAGASIKIAPADAKVRVYITGPPEAQFKAQGRIVYKIKEE 477

Db 12 GATIGKQGHIKQLAREAGASIKIAPREGDVSEWVITGPPEAQFKAQGRIVYKLEE 71

Qy 478 NEVSPKEEVKLEAHIRVPSFAAGRIVGKGKTVNELONLSSAEVVVPDQTPDENQVYV 537

Db 72 NEFPNKEEVKLEAHIRVPSTAGRVGKGTVNELQNLTSAEVIVPROQPDENEDEVIV 131

Qy 538 KITGHFAYACQVAQRKIOEILTOVKQHQ 566

Db 132 RIGHFFASQTQARKLIREIVQVKOQEOK 160

Search completed: April 24, 2003, 16:14:41  
Job time : 24 secs

Result No.	Score	Query Match	Length	DB ID	Description
1	418.5	14.2	854	2 T23837	hypothetical protein
2	243.5	8.2	768	2 T27855	hypothetical protein
3	240	8.1	644	2 A53134	myc far upstream element binding protein
4	136.5	8.0	568	2 T49963	hypothetical protein
5	235.5	8.0	621	2 D96554	hypothetical protein
6	219.5	7.8	313	2 T48439	probable RNA-binding protein
7	229.5	7.8	398	2 T41600	probable pre-mRNA
8	227	7.7	510	2 T38489	onconeural ventral hypothetical protein
9	125.5	7.6	589	2 T19216	hypothetical protein
10	125.5	7.6	611	2 T19217	hypothetical protein
11	222	7.5	479	2 C86225	hnRNP protein - Af
12	219.5	7.4	396	2 S41224	hypothetical protein
13	213.5	7.2	680	2 T25832	alpha-complex protein
14	209.5	7.1	356	2 S58529	high density lipoprotein
15	207.5	7.0	1268	2 A44125	hypothetical protein
16	204.5	6.9	846	2 T04533	hnRNP complex protein
17	204	6.9	413	2 S46109	hnRNP protein E2 - single-stranded nucleic acid binding protein
18	202.5	6.9	365	2 S42471	probable RNA-binding protein
19	202	6.8	362	2 S78515	villin - chicken
20	199	6.8	397	2 T30168	DC stretch-binding domain
21	199.5	6.7	1279	2 T41389	transformation up-regulating protein
22	195	6.6	649	2 E84614	kappa-B motif-binding protein
23	194.5	6.6	1270	2 S23464	protein C08H9.2 [hypothetical protein]
24	190	6.4	463	2 S41495	chicken
25	190	6.4	464	2 S43363	transformation up-regulating protein
26	189	6.4	464	2 T02627	hypothetical protein
27	189.5	6.4	632	2 B88279	protein C08H9.2 [hypothetical protein]
28	184.5	6.2	1198	2 L0PQGTINLRQ	hypothetical protein
29	184.5	6.2	1220	2 T19117	hypothetical protein

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.  
  
OM protein - protein search, using sw model  
Run on: April 12, 2003, 16:11:08 : Search time 21 Seconds  
(without alignments)  
2650.563 Million cell updates/sec  
  
Title: US-09-897-778-176  
Perfect score: 2956  
Sequence: 1 MNKLYIGNLSENAAPSDLES.....VKHQHQOKALQSGPPQSRRK 579  
  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
  
Searched: 283224 seqs, 96134422 residues  
  
Total number of hits satisfying chosen parameters: 283224  
  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries  
  
Database : PR\_73:  
1: pir1:  
2: pir2:  
3: pir3:  
4: pir4:  
  
Listed first 45 summaries  
  
Database : PR\_73:  
1: pir1:  
2: pir2:  
3: pir3:  
4: pir4:  
  
Listed first 45 summaries  
  
Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	418.5	14.2	854	2 T23837	hypothetical protein
2	243.5	8.2	768	2 T27855	hypothetical protein
3	240	8.1	644	2 A53134	myc far upstream element binding protein
4	136.5	8.0	568	2 T49963	hypothetical protein
5	235.5	8.0	621	2 D96554	hypothetical protein
6	219.5	7.8	313	2 T48439	probable RNA-binding protein
7	229.5	7.8	398	2 T41600	probable pre-mRNA
8	227	7.7	510	2 T38489	onconeural ventral hypothetical protein
9	125.5	7.6	589	2 T19216	hypothetical protein
10	125.5	7.6	611	2 T19217	hypothetical protein
11	222	7.5	479	2 C86225	hnRNP protein - Af
12	219.5	7.4	396	2 S41224	hypothetical protein
13	213.5	7.2	680	2 T25832	alpha-complex protein
14	209.5	7.1	356	2 S58529	high density lipoprotein
15	207.5	7.0	1268	2 A44125	hypothetical protein
16	204.5	6.9	846	2 T04533	hnRNP complex protein
17	204	6.9	413	2 S46109	hnRNP protein E2 - single-stranded nucleic acid binding protein
18	202.5	6.9	365	2 S42471	probable RNA-binding protein
19	202	6.8	362	2 S78515	villin - chicken
20	199	6.8	397	2 T30168	DC stretch-binding domain
21	199.5	6.7	1279	2 T41389	transformation up-regulating protein
22	195	6.6	649	2 E84614	kappa-B motif-binding protein
23	194.5	6.6	1270	2 S23464	protein C08H9.2 [hypothetical protein]
24	190	6.4	463	2 S41495	chicken
25	190	6.4	464	2 S43363	transformation up-regulating protein
26	189	6.4	464	2 T02627	hypothetical protein
27	189.5	6.4	632	2 B88279	protein C08H9.2 [hypothetical protein]
28	184.5	6.2	1198	2 L0PQGTINLRQ	hypothetical protein
29	184.5	6.2	1220	2 T19117	hypothetical protein

## ALIGNMENTS

RESULT 1  
T23837  
hypothetical protein M88.5 - Caenorhabditis elegans  
C; Species: Caenorhabditis elegans  
C; Date: 15-Oct-1999 #sequence\_change 15-Oct-1999  
C; Accession: T23837  
R; Submitted to the EMBL Data Library, June 1994  
A; Reference number: Z19806  
A; Accession: T23837  
A; Status: preliminary; translated from GB/EMBL/DDBJ  
A; Molecule type: DNA  
A; Residues: 1-894 <WILD>  
A; Cross-references: EMBL:Z34802; PIDN:CAA84338.1; GSPDB:GN00021; CESP:M88.5  
A; Experimental source: clone M88  
A; Genes: CESP:M88.5  
A; Map position: 3  
A; Introns: 38/3; 96/1; 133/3; 255/3; 307/3; 395/3; 436/2; 492/3; 657/3; 759/1; 820/1

Query Match 14.28; Score 418.5; DB 2; Length 854;  
Best Local Similarity 24.1%; Pred. No. 6-19;  
Matches 160; Conservative 103; Mismatches 215; Indels 187; Gaps 27;

Qy 78 QRIKQIQRNIPPHLOWEVLDSLLVQYGVESCVNTSETAVN-----VYSSK 129  
Db 167 QMQQQOSNSOASAHEMHHQOLQAOVQQ----QAQMHHRLQGAPINPQQFMVPPTMMQP 221  
Qy 130 DOARQALDKLNGFQLENFTLKRYATPDEMA-AQO--NPLQQPRGRGLGQRGSRSRQGSP 185  
Db 222 QMQQA-QQQQAOQMHQMIHQHQHOMQHQAHQHQHOMQHQHQHOMQHQHQHOMQHQHQ 279  
Qy 186 G-----SVSKQKP-----C---DLPPLRVYPQFYGAIGIKEGATIRNITPQ 224  
Db 280 NHNQRHNQHSISGHPHIPQNLMPPCMKLWPKIROVGEVKHYAVTIGPNGSTKDIASS 339  
Qy 225 TOSKID----VHKRKENA-GAAEKSTIILSPEGTSAACKSILEIMKHEA-QDIFKTEEL 277  
Db 340 TCRDVFLVNLSKERTVLGNNDRLTVHGAEQATAKVARILDVQSEAYKDDNVGADT 399  
Qy 278 PIKILAINNFYVRLIGKEGRNLKKIQQDODTKITISPLQE-----LT-----LYNPERT 326  
Db 400 VLRMRANQLCRLIGKAGSSKIEQKGTGNNITYKVEPPGGISGLTANELLGMLERT 459  
Qy 327 ITVKNG-NVETCAKAERIMKIRESYENDTASMLDAHLTPGLNINALGLEPPSGMPP- 384  
Db 460 IMVRGPSEAVVQAELALISAKKKCYESD-SQLRAQSMQP-----MPMM-MPPI 508  
Qy 385 -PTSGPPSAMTPY-----POEOFSETETVHO----- 410  
Db 509 LPPGASSSAVASAPHFTPTPVGYMQVQHFASSOHLVHQNANNNSFLQFGVILQIOPGTTNLRQ 568

Qy	411	- - - FIPALSYGAIIGKOGQHIKOLSRPAGASIKIAPAEPAK - - - - -	451	Qy	441	IAPAEAPDAKVRMVIITGPPEAQEKAQGRIVGKIKE- - - ENFVSPKKEEVLEA - - - - -	490
Db	569	VRMVYPSNIGALIAGKGNIKMILIRDGAVKI - - - EAPEEKTOQEAEAKKKERKLDER	625	Db	486	EKPDTNPNSEDTAVINGTRDIOYRATERITEIVARAIKINGAPDRGSAGTVLPQGQSF	545
Qy	452	- - - - - RMYITITGPPEAQEKAQGRIVGK - - - - -	473	Qy	491	HIVYPSFAAGRYTGKGKTVNELQNLQSAEV- VVPRDQTDPENDOYVKITGH - - - EYA	545
Db	626	DSGCEGVAASGDHPQEFLEDNATINSSDAIEEKPKPVSERWYTINGDDLQLIKAQSYVFSK	685	Db	546	YMVPAZCGLYVKGGENIKOIERETGATGLAPAEQKNE- DEKFELKGSQLOIRHA	604
Qy	474	IKEENFVSP- - - - - KEEVKDEAHTRVPSFAAGRYGKRTYNEQNLQSAEVVVRP	525	Qy	546	CQVAQRKIQEILITQVK - - - - - QHQQKALQSPQPS	576
Db	686	IAETSSLSSSGMDGDRSHMLRIRTEVSPTRIGRIGKGSQNVRELQRTGAVKIKE	745	Db	605	SHLVRIKYGELISPNTPVPPLOGAGGSYQQQAMEFSAGTQN	644
Qy	526	DQTPP- - - - - ENDQVVKITGHYACOQAQRKIQEILITQVKO- - - HQQOKA-LQS	571				
Db	746	BERRGEVYRHDGQLEEMDTMIRTIGNMYSTHNVQFLAHIVNEYRSGDHRNKSSDYKG	805		RESULT 3		
Qy	572	GPPQS	576		A53184		
Db	806	GRPHS	810		N; Alternative names: FUSE-binding protein		
				C; Species: Homo sapiens (man)			
				C; Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999			
				C; Accession: A53184			
				R; Duncan, R.; Bazai, L.; Michelotti, G.; Tomonaga, T.; Krutzsch, H.; Avigan, M.; Leve			
				Genes Dev. 8: 465-480, 1994			
				A; Title: A sequence-specific, single-strand binding protein activates the far upstream			
				A; Reference number: A53184; PMID:94170991; PMID:8125259			
				A; Status: Preliminary			
				A; Molecule type: mRNA			
				A; Residues: 1-644 <DUN>			
				A; Cross-references: GB:U05040; NID:9460151; PIDN:AAA17976_1; PID:9460152			
				C; Keywords: DNA binding			
				Query Match	8.1%	Score 240;	DB 2;
				Best Local Similarity	22.3%	Pred No. 1, 2e-07;	
				Matches 100;	Conservative	Mismatches 154;	Indels 120;
						Gaps 17;	
				Qy	161	QONPLQ- - - QPRGRRLGQRGSSRQGSPGSVKQPKDPLPLRLLVPTQFYAGTIGKEGAT	217
				Db	63	QKRPLEDDGQDPAKVAQNDSDGTQJLPPMIQOQQSSVMTEYKVPDGMWGFIRGGEQ	122
				Qy	218	IRNITKOTOSKIDVHRKENAGAAEKSSTTISPECSAAACKSITLEJMHRKAQDKIFTE- -	275
				Db	123	123 ISRIQEQSCSKQI- APPGGGLPERSMQMLTGPESVQSAKRLDQIVEKRPAPCPFHGD	181
				Qy	276	-- - EIPKLTLAHNNFVGRLIGKEGRNLKKIEODTDKTTISPLQELTYQPERT- - - I 327	
				Db	182	GPGNAVQELIMIPASKAGAIVVKGGETIKQLQRAYGVKYM - IQD - - - CPONTGADKPL	235
				Qy	328	TVKGNVETCAKAEELIMKKIRE- - - SYENDIASMNQAHLPGLNLNGLFPPTSGM	382
				Db	236	RITEDPYKVQQAKEMVLLIRDQGGFREVNEYGSR- - - - - IGG - - - - - NEGI	278
				Qy	383	PPPTSGPPSAMTPPYPPQESESETVQFIPALSVGAIIKGQKHIKLSRFAGASIKTA	442
				Db	279	DVP- - - - -	
				Qy	443	PAAEPDAKVRMVIITGPPE-AOFKA	473
				Db	313	PDGTTPE- RIAQITGPDRQCHAETITDLRSVQAGNPGGPGCPGRGRRGQGNWNMG	371
				Qy	474	-- - IKEENFVSPKEEVKLEAHITRVPSFAAGRYVKGKTYNELQNLSSAEVYVPRDQTP	529
				Db	372	PPGGLQEFNF- - - - - - - - - VPGKTGLLIGKGGETIKSISQSGARIELQRNPPP	418
				Qy	530	DENDQV-VVKITGH- - - FYACQVAQRKI	553
				Db	419	NADPNMKLFETIRGTPQDIDYARQLEEKI	447
				RESULT 4			
				T49952			
				C; Species: Arabidopsis thaliana (mouse-ear cress)			
				C; Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000			

C:Accession: T49622 R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le submitted to the Protein Sequence Database, April 2000	Query Match 8.0%; Score 235.5; DB 2; Length 621; Best Local Similarity 22.9%; Pred. No. 2.2e-07; Matches 104; Conservative 76; Mismatches 171; Indels 103; Gaps 17;
A:Reference number: 224493	Qy 199 RLLVPQFGAGTIRNITKOTOSKI-----D 230 Db 22 RLCPATRGTAIGGGVIRHLQSVIGSKTRVIDIPPEERVVLLAAPSCKKKDESN 81
A:Status: preliminary	Qy 231 VHRKENAGA---AEKSTILS-----PEGSTAAKSYLEIMHK----PQDI-----271
A:Molecule type: DNA	Db 82 VCDSEBPGSDPKDBKGSCAGTGGDDEAPSSAQMLLRRVERIEGDDATVDGDEL 141
A:Residues: 1-568 <BEV>	Qy 272 -KTEFEIPLKLHANNFVGRLIGKEGRNLKJEDTDKUTISPLQEL--TLYNPERTIT 328
A:Cross-references: EMBL:AL353993; GSPDB:GN00063; ATSP:F8M21.160	Db 142 DKGESEGLCMTIVGNQDLMASKGMQIRESDGAIVRSSTDQLPPCAFPGDVVHQ 201
A:Experimental source: cultivar Columbia; BAC clone F8M21	Qy 329 VKGNVETCAKAEEFIMMKIRES-----YENDIASMNLQAHLPGLUNLNAL 373
A:Genetics:	Db 202 MNKFSESVKIALLLVLTNCLEQESSAPPWDECPPCPGPYPPYEISME-----247
A:Gene: ATSP:F8M21.160	Qy 374 GLFPPTSGMPPTSGPPSAMTPPYPOFEQESTETVHOFIPALSYGATIGKOGQHITQLSR 433
A:Map position: 5	Db 248 -YHPWDHPNPMPED--VGPFNRRVEEVAVRLLPADKVGSLGKGAIVRALQN 303
A:Introns: 200/3; 337/3; 544/3	Qy 434 FAGASIKIAEADPAKAYRMVITLGPPPEAQKF---AQG---RIYGKIREENVSPKEVK 487
Query Match 8.0%; Score 236.5; DB 2; Length 568;	Db 304 ESGASITKVS-D-PTHSERLIVISARENLERHSLAQDGDMYRHNRIVEGF--EPSAA 359
Best Local Similarity 22.8%; Pred. No. 1.7e-07;	Qy 488 LEAHTRVPSFAAGRVIGKGKTYNELONLSSAAEV-VVVKITFGHFEV 544
Matches 92; Conservative 79; Mismatches 144; Indels 89; Gaps 16;	Db 360 VVARLLHSPVIGRLLGGHLISEMRATASIRVFAKDQATKYESODEIVQVGNL- 418
Qy 195 DPLRLVLPQTGVGAIIGKEGATIRNITKOTOSKIDVHRRENAGAAEKSTILSPEST 254	Qy 545 ACQVAKRQIOELITCOKHOQOKAR---QGPP 574
Db 49 DTVERLYCPVKKIGSVIGGGDTVKOLNDTSKIRI-GEAIPGCDERVITYSPSDET 107	Db 419 -KTVQDALFQILCRLRAFMPGRLPPQGMGPP 450
Qy 255 A-----ACKSILEIMHKAEQDITPTEETP-----LKLIAHNNFVGRJLKGKR 297	...RESULT 6
Db 108 AFGDGEKVLSPAQDAFLFRTHVVADDERSDSEPEGEKQVTAKLVPSDQIGCLGLRGQ 167	t4639 probable RNA-binding protein - Arabidopsis thaliana
Qy 298 NLKRIEQDPIKTKITISPLQBLTL--YNPERITIVKGNAVTCOKABEINMKIRESYND 355	N; Alternative names: Protein T22M21_30
Db 168 IVQNIRSETGAQIRTYKDRMPLCALNSDELIOISGEVLTVKKALLQIASRL--HENPS 224	C; Species: Arabidopsis thaliana (mouse-ear cress)
Qy 356 JASHNLQAHPLGLNLAQNLGLFPPPS-----GMP-----PTSCP 390	C; Date: 20-Apr-2000 #sequence_revision 20-Apr-2000
Db 225 IRSQNL-----LSSGGYPASLMSHAGGPRLVGLAPLGMGSGRDQGDWSRPLQP 275	C; Accession: T48439
Qy 391 SAMTTPPFQFQESETETVHDFIPALSYGAIIGKOGHIIKOLSFAGASIKI--APAEPAD 448	R; Bevan, M.; Terlyn, N.; Ardiles, W.; Buyshaert, C.; Dasseville, R.; De Clerck, R.; ewes, H.W.; Rudd, S.; Lenicke, K.; Mayer, K.F.X.
Db 276 -RNDPATEF-----FIRLVSPLYNTIASIVKGGAATNLQERETRATLKVDSSRTEGND 328	Submitted to the Protein Sequence Database, March 2000
Qy 449 AKYRNVVITGPP-----EAOFKAQGRYKGKIREENFSPEEVKLEAHTRVPSFA 498	A; Reference number: Z24487
Db 329 C---LITISAREVFEDAYSPTEAVMRLOPKCSDKVDRDSL-----VSFTTRLLVPSR 380	A; Status: preliminary
Qy 493 AGRVIGKGKTYNELONLSSAAEV-VVPRDQTP---DENQVVV 537	A; Molecule type: DNA
Db 381 TGCLGKGAITTEMRTANIRTLGENLPKVASDDENYQV 424	A; Residues: 1-313 <BEV>
...RESULT 5	A; Cross-references: EMBL:AL162875
D96554 hypothetical protein F19C24.19 [imported] - Arabidopsis thaliana	A; Experimental source: cultivar Columbia; BAC clone T32M21
C:Species: Arabidopsis thaliana (mouse-ear cress)	C; GenBank: T32M21_30
C:Accession: D96554	C; Date: 20-Apr-2000 #sequence_revision 02-Mar-2001
R:Theoretical, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chung, M.K.; Conn, L.; Conway, A.B.; Venter, T.H.; Dewar, K.; ansen, N.; Hughes, B.; Hvizdar, L.	C; Genetics:
N:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Schwartz, J.R.; Shiann, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.	A; Map position: 5
A:Title: Sequence and analysis of chromosome 1 of the Plant Arabidopsis.	A; Intron: 29/1; 108/3; 144/2; 209/1; 229/3; 261/3; 282/2
A:Reference number: A66141; PMID:21016719	A; Note: T32M21_30
A:Accession: D96554	Qy 183 GSPGSVSKQKP-----CDLP--LRLVPTQEVGAIGKEGATIRNITKOTOSKID 230
A:Molecule type: DNA	Db 10 GSPEELAKRSPHDSEASAKRSPHTRFVLSNAAGSYIGKGSTTTEFOQAKSGARIQ 69
A:Residues: 1-601 <STO>	Qy 231 VHRKEN--AGAAEKSITLSTPEGSTAAKSITLEIMHKAEQDIKTFEELP--LKLIAHN 285
A:Cross-references: GB:AE005173; NID:911094762; PIDN:AG29695.1; GSPDB:GN00141	Db 70 LSRNQEFPGSTDRIMISSLSSIKEVNGNLEFLDKLHSLEHAEDNEVRPRRIRLVPN 129
A:Gen: F19C24.19	Qy 286 NFVGRLLGKEGRNLKKEQDTDKUTISPLQELTN-PERTITVKGNTVETAKAEELIM 344
A:Map Position: 1	

Db	130	SSCGGIGKGGATISFIESKAGIKSPLDN-TFGLSLRVLTLGGFFEEQMRAIDLIL 188	A; Reference number: I38489; MUID: 94000830; PMID: 8398153 A; Accession: I38489 A; Status: Preliminary; translated from GB/EMBL/DBJ A; Molecule type: mRNA A; Residues: 1-510 <RES> A; Cross-references : EMBL:U04840; NID:9440877; PIDN:AAA16022.1; PID:9440878 C; Genetics: A; Gene: Nova-1
Db	345	KKTRESEYNDIASKMLQAHLIPGLNLNALGLEPPSGMPPPTSGPSSAMTPYPQQESE 404	Query Match Score 227; DB 2; Length 510; Best Local Similarity 22.4%; Pred. No. 5.8e-07; Matches 119; Conservative 77; Mismatches 175; Indels 156; Gaps 20;
Db	189	AKIT--EDDIYSONWPSYASAGYVNPAGNS-- 459	A; Reference number: I38489 A; Accession: I38489 A; Status: Preliminary; translated from GB/EMBL/DBJ A; Molecule type: mRNA A; Residues: 1-510 <RES> A; Cross-references : EMBL:U04840; NID:9440877; PIDN:AAA16022.1; PID:9440878 C; Genetics: A; Gene: Nova-1
Db	233	STTVTGADBHIGLVLGRGRNIMEIQMTGARIKISDRGDFMSGTDRKV\$--ITGP 289	Query Match Score 227; DB 2; Length 510; Best Local Similarity 22.4%; Pred. No. 5.8e-07; Matches 119; Conservative 77; Mismatches 175; Indels 156; Gaps 20;
Db	460	PEAQKAGQRYGKI 474	A; Reference number: I38489 A; Accession: I38489 A; Status: Preliminary; translated from GB/EMBL/DBJ A; Molecule type: DNA A; Residues: 1-510 <RES> A; Cross-references : EMBL:U04840; NID:9440877; PIDN:AAA16022.1; PID:9440878 C; Genetics: A; Gene: Nova-1
Db	290	QRAIQQAETMIKQV 304	Query Match Score 227; DB 2; Length 510; Best Local Similarity 22.4%; Pred. No. 5.8e-07; Matches 119; Conservative 77; Mismatches 175; Indels 156; Gaps 20;
RESULT 7	T41650	probable pre-mRNA splicing factor - fission yeast (Schizosaccharomyces pombe)	Query Match Score 227; DB 2; Length 510; Best Local Similarity 22.4%; Pred. No. 5.8e-07; Matches 119; Conservative 77; Mismatches 175; Indels 156; Gaps 20;
C; Date: 01-Dec-1999 #text_change 03-Dec-1999	C; Accession: T41650	Schizosaccharomyces pombe	Query Match Score 227; DB 2; Length 510; Best Local Similarity 22.4%; Pred. No. 5.8e-07; Matches 119; Conservative 77; Mismatches 175; Indels 156; Gaps 20;
R; Lyne, M.; Radford, M.A.; Barrell, B.G.; Bothé, G.; Pohl, T.	A; Reference number: T41650	submitted to the EMBL Data Library, October 1998	Query Match Score 227; DB 2; Length 510; Best Local Similarity 22.4%; Pred. No. 5.8e-07; Matches 119; Conservative 77; Mismatches 175; Indels 156; Gaps 20;
A; Experimental source: strain 972h-; cosmid C757	A; Accession: T41650	A; Status: Preliminary; translated from GB/EMBL/DBJ	Query Match Score 227; DB 2; Length 510; Best Local Similarity 22.4%; Pred. No. 5.8e-07; Matches 119; Conservative 77; Mismatches 175; Indels 156; Gaps 20;
A; Residues: 1-398 <LEN>	A; Accession: AL031825; PIDN:CAA21234.1; SPDB:GN000668; SPDB:SPCC757.09C	A; Molecule type: DNA	Query Match Score 227; DB 2; Length 510; Best Local Similarity 22.4%; Pred. No. 5.8e-07; Matches 119; Conservative 77; Mismatches 175; Indels 156; Gaps 20;
A; Cross-references : EMBL:AL031825; PIDN:CAA21234.1; SPDB:GN000668; SPDB:SPCC757.09C	A; Experimental source: strain 972h-; cosmid C757	A; Cross-references : EMBL:AL031825; PIDN:CAA21234.1; SPDB:GN000668; SPDB:SPCC757.09C	Query Match Score 227; DB 2; Length 510; Best Local Similarity 22.4%; Pred. No. 5.8e-07; Matches 119; Conservative 77; Mismatches 175; Indels 156; Gaps 20;
C; Genetics: A; Gene: SPDB:SPCC757.09C	A; Accession: T41650	A; Molecule type: DNA	Query Match Score 227; DB 2; Length 510; Best Local Similarity 22.4%; Pred. No. 5.8e-07; Matches 119; Conservative 77; Mismatches 175; Indels 156; Gaps 20;
A; Map position: 3	A; Map position: 3	A; Map position: 3	Query Match Score 227; DB 2; Length 510; Best Local Similarity 22.4%; Pred. No. 5.8e-07; Matches 119; Conservative 77; Mismatches 175; Indels 156; Gaps 20;
Db	176	DRESSRGSPGS--VSXGKPCD--LPIRLULYQFGAIKGKEGATIRNITK 223	Query Match Score 229.5; DB 2; Length 398; Best Local Similarity 26.5%; Pred. No. 2.8e-07; Matches 90; Conservative 55; Mismatches 148; Indels 47; Gaps 11;
Db	62	DQSMNTPQEPISQVPPISAKPMDVATQQLTLALLSSTREGTIGKAGKNAELRS 121	Query Match Score 229.5; DB 2; Length 398; Best Local Similarity 26.5%; Pred. No. 2.8e-07; Matches 90; Conservative 55; Mismatches 148; Indels 47; Gaps 11;
Db	224	OTOSKIDYHRKENAGAAEKSITILSTPEGTSACKSLEIMMHKAQDQDKFTEE---IP 278	Query Match Score 229.5; DB 2; Length 398; Best Local Similarity 26.5%; Pred. No. 2.8e-07; Matches 90; Conservative 55; Mismatches 148; Indels 47; Gaps 11;
Db	122	FTRVKARY-TRAVPNHDRVLTISGPLEVNVRAYRTIDFAKNTNDGTSANDPRK 189	Query Match Score 229.5; DB 2; Length 398; Best Local Similarity 26.5%; Pred. No. 2.8e-07; Matches 90; Conservative 55; Mismatches 148; Indels 47; Gaps 11;
Db	279	LK1LAHNNFVGRLKQKTEQDDTKTISPLQEYLNPERTIVKGNEVETAK 338	Query Match Score 229.5; DB 2; Length 398; Best Local Similarity 26.5%; Pred. No. 2.8e-07; Matches 90; Conservative 55; Mismatches 148; Indels 47; Gaps 11;
Db	181	LRLLIAISLMGSIIGNGLRKLQDKSCRMIAS-KDMLPQSTERVEIIGTVNDLHA 238	Query Match Score 229.5; DB 2; Length 398; Best Local Similarity 26.5%; Pred. No. 2.8e-07; Matches 90; Conservative 55; Mismatches 148; Indels 47; Gaps 11;
Db	339	AEBBIRKIREYNDIASN-NIQLAH--IPLGLNLNA-LGLFPPTSGMPPTSGP-- 389	Query Match Score 229.5; DB 2; Length 398; Best Local Similarity 26.5%; Pred. No. 2.8e-07; Matches 90; Conservative 55; Mismatches 148; Indels 47; Gaps 11;
Db	239	AIWEIGCLIDDWERGACTIVFNVSRLTOPLSLASTASPQVSPAA-ESTTSEAI 296	Query Match Score 229.5; DB 2; Length 398; Best Local Similarity 26.5%; Pred. No. 2.8e-07; Matches 90; Conservative 55; Mismatches 148; Indels 47; Gaps 11;
Db	390	-----PSMTAPPYQFQEQSETETVHQFLPSGATIGKQGHKOLSRFGA 437	Query Match Score 229.5; DB 2; Length 398; Best Local Similarity 26.5%; Pred. No. 2.8e-07; Matches 90; Conservative 55; Mismatches 148; Indels 47; Gaps 11;
Db	297	PENFSYGAQVEPATOMP---FLQOPKVTONISTADMYGCTIGRRGSKKISEIRTSGS 352	Query Match Score 229.5; DB 2; Length 398; Best Local Similarity 26.5%; Pred. No. 2.8e-07; Matches 90; Conservative 55; Mismatches 148; Indels 47; Gaps 11;
Db	438	SIIKIAAREAPDAKVRMIVITGPPEAQAFQAGRIYGIKKE 477	Query Match Score 229.5; DB 2; Length 398; Best Local Similarity 26.5%; Pred. No. 2.8e-07; Matches 90; Conservative 55; Mismatches 148; Indels 47; Gaps 11;
Db	353	KISTAKEPHDETGERMFITGTHEENERALKFLYQIEME 392	Query Match Score 229.5; DB 2; Length 398; Best Local Similarity 26.5%; Pred. No. 2.8e-07; Matches 90; Conservative 55; Mismatches 148; Indels 47; Gaps 11;
RESULT 8	I38489	concurrent ventral antigen-1 - human	Query Match Score 225.5; DB 2; Length 589; Best Local Similarity 23.4%; Pred. No. 3e-07; Matches 86; Conservative 72; Mismatches 107; Indels 53; Gaps 11;
C; Species: Homo sapiens (man)	C; Accession: I38489	C; Species: Homo sapiens (man)	C; Genetics: A; Gene: CESP:C12D8-1a A; Map position: 5
C; Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 05-Nov-1999	C; Accession: I38489	C; Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 05-Nov-1999	C; Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 C; Accession: T19216 A; Status: Preliminary; translated from GB/EMBL/DBJ A; Molecule type: DNA A; Residues: 1-589 <RES> A; Cross-references : EMBL:Z13969; PIDN:CAA98232.1; GSPDB:GN00023; CESP:C12D8.1a A; Experimental source: clone C12D8 C; Genetics: A; Gene: CESP:C12D8-1a A; Map position: 5
Db	239	AIWEIGCLIDDWERGACTIVFNVSRLTOPLSLASTASPQVSPAA-ESTTSEAI 296	Query Match Score 225.5; DB 2; Length 589; Best Local Similarity 23.4%; Pred. No. 3e-07; Matches 86; Conservative 72; Mismatches 107; Indels 53; Gaps 11;
Db	390	-----PSMTAPPYQFQEQSETETVHQFLPSGATIGKQGHKOLSRFGA 437	Query Match Score 225.5; DB 2; Length 589; Best Local Similarity 23.4%; Pred. No. 3e-07; Matches 86; Conservative 72; Mismatches 107; Indels 53; Gaps 11;
Db	297	PENFSYGAQVEPATOMP---FLQOPKVTONISTADMYGCTIGRRGSKKISEIRTSGS 352	Query Match Score 225.5; DB 2; Length 589; Best Local Similarity 23.4%; Pred. No. 3e-07; Matches 86; Conservative 72; Mismatches 107; Indels 53; Gaps 11;
Db	438	SIIKIAAREAPDAKVRMIVITGPPEAQAFQAGRIYGIKKE 477	Query Match Score 225.5; DB 2; Length 589; Best Local Similarity 23.4%; Pred. No. 3e-07; Matches 86; Conservative 72; Mismatches 107; Indels 53; Gaps 11;
Db	353	KISTAKEPHDETGERMFITGTHEENERALKFLYQIEME 392	Query Match Score 225.5; DB 2; Length 589; Best Local Similarity 23.4%; Pred. No. 3e-07; Matches 86; Conservative 72; Mismatches 107; Indels 53; Gaps 11;

Qy	202	VPTQFVGAIIGKEGATINNIKQTQSKIDIVRKENAGAAEKSITILSTPEGTSAACKSIL	261	Qy	490	A--HIRPSSFAAGRIVTGKGKTVMELQNLSSAEVYVPRDQTPDENQVYVKITGHFYAC	546
Db	54	IPESAVGIVIGRSSEIICAGKACRQMSPDAPSSGVRMTELEGSRNVETAKHLIN	113	Db	339	ASTYMSPAAKCGLVIGKGGETIKQINSESGAHELSRDPGNADEKVFV-IKGKRAI	397
Qy	262	EIMKHEADQIKFTEBIP---LKLIAHNEVGRGLIGKGRNKKIEQDTK-ITISPL	315	Qy	547	QVAQRKIQ 554	
Db	114	EVVAR-SQNERPQVGFPAQTTIDALPPNRCGLTIGKSDTIDROQEKSCKMILVQDN	172	Db	398	EHAKHLIR 405	
Qy	316	QELTLNPERITVKGNVETCAKAEELMKKIRESYENDIASMNLQAHLLPOLNLNALGL	375	RESULT 11			
Db	173	QSVS--DOSKPLRTGDPKIELA-KOLVAAILNSGGDNGSCLQMEHAGG-----	221	Q86275	hypothetical protein [Imported] - Arabidopsis thaliana		
Qy	376	FPTTSMMPPTSGPPSAMPTPPQFEQSETETVHQFLPALSVALIGKQGOHIIKLSREA	435	C;Species: Arabidopsis thaliana (mouse-ear cress)			
Db	221	-----GGASA-----RGEVV--VPRSSVGIIGKGQDITKRALMET	256	C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-dec-2001			
Qy	436	GASIKTAPAPA PDAKVRWVITGSPPEAQKFAQGRIYKGKKEENFVS----PKEEVKLE	489	R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, Chin, N.F.; Hughes, B.; Huizar, L.			
Db	257	GRKQFKPDDDPSPERCAVIMTDRQYATERTELYKSTMQGGGGNAGAMVSNE	316	Nature 408, 816-820, 2000			
Qy	490	A--HIRPSSFAAGRIVTGKGKTVMELQNLSSAEVYVPRDQTPDENQVYVKITGHFYAC	546	A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marzia Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.			
Db	317	ASTYMSPAAKCGLVIGKGGETIKQINSESGAHELSRDPGNADEKVFV-IKGKRAI	375	A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.			
Qy	547	QVAQRKIQ 554		A;Title: Sequence and analysis of chromosome 1 of the Plant Arabidopsis.			
Db	376	EHAKHLIR 383		A;Reference number: A86141; MUID:21016719; PMID:11130712			
Qy	1917	T1917		A;Accession: c86275			
Db				A;Status: preliminary			
Db				A;Molecule type: DNA			
Db				A;Residues: 1-479 <STOP>			
Db				A;Cross-references: GB:AE005172; NID:95080792; PIDN:AD39302.1; GSPDB:GN00141			
Db				C;Genetics:			
Db				A;Map position: 1			
Qy	202	VPTQFVGAIIGKEGATINNIKQTQSKIDIVRKENAGAAEKSITILSTPEGTSAACKSIL	261	Query Match Score 7.5%; Length 479;			
Db	136	EVVAR-SNPRFQYGPRAOTTIDAPPNRCOLIGKSGTIDROQEKSCKMILVQDN	194	Best Local Similarity 21.6%; Pred. No. 1..e-06;			
Qy	316	QELTLNPERITVKGNVETCAKAEELMKKIRESYENDIASMNLQAHLLPOLNLNALGL	375	Matches 98; Conservative 93; Mismatches 169; Indels 94; Gaps 19;			
Db	195	QSVS--DOSKPLRTGDPKIELA-KOLVAAILNSGGDNGSCLQMEHAGG-----	243	Db			
Qy	262	EIMKHEADQIKFTEP---LKLIAHNEVGRGLIGKGRNKKIEQDTK-ITISPL	315	165 LQOPRORG ----LGQGSSPQGSPPGSYSKQKPC -DLPLRILVPTOFQYGAIIKGKEA 216			
Db	136	EVVAR-SNPRFQYGPRAOTTIDAPPNRCOLIGKSGTIDROQEKSCKMILVQDN	194	Db	5 LRNIHGKRSNLSQSEFTGNGGSRRNHDPEQNTASEPDVRYLCPVKRKGTSIGKGGE 64		
Qy	316	QELTLNPERITVKGNVETCAKAEELMKKIRESYENDIASMNLQAHLLPOLNLNALGL	375	Qy	217 TIRNITKOTOSKIDIVRKRENAGAAEKSITILSTPEGTG -----ACKSTLEIHKR -- 266		
Db	195	QSVS--DOSKPLRTGDPKIELA-KOLVAAILNSGGDNGSCLQMEHAGG-----	243	Db	65 IAKQTRSETSNMRIN-BALPGCEERVYTMSTBNBNEHGDGELVCPA-LDALEKVDH 122		
Qy	202	VPTQFVGAIIGKEGATINNIKQTQSKIDIVRKENAGAAEKSITILSTPEGTSAACKSIL	261	Qy	267 -----BAQDKTFTEPIKILAHNFVGRGLIGKGRNLKKEQDTDTKTI - 312		
Db	136	EVVAR-SNPRFQYGPRAOTTIDAPPNRCOLIGKSGTIDROQEKSCKMILVQDN	194	Db	123 MVVADADDQDGTDDNDGEKQFTVNMVLPSDQIGCVKGKGGVIONLNDTNQIRV 182		
Qy	316	QELTLNPERITVKGNVETCAKAEELMKKIRESYENDIASMNLQAHLLPOLNLNALGL	375	Qy	313 ---SPLQELTLYNPERTITVKGNVETCAKAEELMKKIRESYENDIASMNLQAHLLPOLNLNALGL 366		
Db	195	QSVS--DOSKPLRTGDPKIELA-KOLVAAILNSGGDNGSCLQMEHAGG-----	243	Db	183 KDHPACALTSHDPLLTI -GEPVVREALYQVASLLDNPSPRFQHLLSSSSMMOP 241		
Qy	262	EIMKHEADQIKFTEP---LKLIAHNEVGRGLIGKGRNKKIEQDTK-ITISPL	315	Qy	367 GLNLNALGFLRPPTSEMPPTSGPSPSAMTPPYPOFE-QSETETVHQF ----IPALSVGA 420		
Db	136	EVVAR-SNPRFQYGPRAOTTIDAPPNRCOLIGKSGTIDROQEKSCKMILVQDN	194	Db	242 GAMLM-----AALTSSHRRNAYVRRLADAREFCYCFCPAENVGGV 283		
Qy	316	QELTLNPERITVKGNVETCAKAEELMKKIRESYENDIASMNLQAHLLPOLNLNALGL	375	Qy	421 IGKOGHIIKOLSRPAGASKIAPEADDAKTYMVITGPE ---AQFRQGRI 470		
Db	195	QSVS--DOSKPLRTGDPKIELA-KOLVAAILNSGGDNGSCLQMEHAGG-----	243	Db	284 IGKGGFTRNQPRQETGATIRVNTESETDDD - CLIFISKEKFEDQSPAVNAATLQQR 341		
Qy	376	FPTTSMMPPTSGPPSAMTPPYPOFEQSETETVHQFLPALSVALIGKQGOHIIKQLSRA	435	Qy	471 VCKI-KEENPVSPKEVKILEAHTRVPFAAQRVIGKGKTVMELQNLSSAEV-VVPRDQT 528		
Db	244	-----GGASA-----RGEVV--VPRSSVGIIGKGQDITKRALMET	278	Db	342 SKVGRDN-----DIASTRILLVSSQICLIRKGAVISENSVPRANRILQKEDV 395		
Qy	436	GASIKTAPAPA EADAKYRMVITGPDEAQAFQGRYKIKEENFVS----PKEEVKLE	489	Qy	529 PD -ENPQVYKTFCHYEQVYQQRKIQEIUTQ 560		
Db	279	GTKIQFKPDDPSTPERCAVIMGTRDQYRATERITELVKSTMOQGGGGNVAGAMVSNE	338	Db	396 PKTARDEEWMQITGSPDAMKA-----LTQV 422		
				RESULT 12			

S41224	African clawed frog C:Species Xenopus laevis (African clawed frog)	Best Local Similarity 21.3%; Pred. No. 6.5e-06; Matches 73; Conservative 69; Mismatches 139; Indels 61; Gaps 9;
C:Accession: SA1224	C:Accession: SA1224	QY 202 VPTQFQAIIGKEGATIRNITKOTOSKIDVHRENAGAAEKSTILSTPGETSAACKSIL 261.
R:Siem, H.; Matunis, M.J.; Michael, W.M.; Dreyfuss, G.	Nucleic Acids Res. 21, 1191-1198, 1993	Db 152 TBESEGLVGRNGEIVTQI:QISNSQRVQI-VAEPISTGYRDIYSENIEVAKRLN 210
A:Title: The pre-mRNA binding K protein contains a novel evolutionarily conserved motif.	A:Reference number: S41224; PMID:93219080; MUID:4464704	Db 262 EIM--HKEAQD-----IKFTEIPKILAHNNVGRGLIGKEGRNLKKIEQ 304.
A:Accession: S41224	A:Molecule type: mRNA	Db 211 EVVARGRKLKSQEPLPCSVEOFPIAVSNSSKVTIIPANCGATGKGEQRKLRS 270
A:Residues: 1-396 <S10>	C:Superfamily: transformation upregulated nuclear protein hnRNP K	Db 305 DNDTKITISPLQETLYNPERITVKGVRCAKEEEIMKKRESYENDIASMNLQAH 364.
Query Match Score 7.48; Length 396;	Best Local Similarity 21.3%; Pred. No. 1.2e-06; Matches 93; Conservative 76; Mismatches 166; Indels 101; Gaps 18;	Db 271 WNCDFL--IQENNTIAVKPIQITSGPKVEHHKA-----LVADY 310
Db 17 GQRSSRQSGPSVKOKPCDL--PLRLVTPQFVGATIGREGATIRNITKOTOSKIDVH 232.	Db 365 IPGLNUNALSLIFPPTSGMPPPSAMTPYPOEQSETETVHQFIPALSGVALIGKQ 424	
Db 17 GKRPAEDMEQAFKSRNTDMVFLRILQSNAGAVIGKGGKNIKALRTDYASVSV- 75.	Db 311 LDGFD-----ECPPAGMAGNSPVAMS---LQVKVPRSTVGAIMGLQ 349	
QY 233 RKENGAAREKATIILSTPEGTTAACKSILEMIMAEOADIKPTEIPLKILAHNNFVGRLI 292	Db 271 WINCDFL--IQENNTIAVKPIQITSGPKVEHHKA-----LVADY 310	
Db 76 -PSSGPERILSISADIEETGEILKKIPLTEHFGNDF--DCELRLLIHQSLAGGT 131	Db 350 GSNIKAISETETKIQFNPDDKMLERTLVGAKRNKVYVC ALRQKIVEANSNEANT 408	
QY 293 GKEGRNLKKEIQQDTDKITISPLQETLYNPERITYGVNEYTCAKAEEBIMMKRIES- 350	Db 485 EYKLEAHIRVPSEAGRYIGKGGKTYNELQNLSSSEVVYVPRD 526	
Db 132 GVKAKIKELREKTQ-T-TIKLFQCCPHISTDRVLIGGRPRVYCEKIVLDLISESPV 189	Db 409 PISL-FYMLIPASRCGVLVGRGETEIQINKESGAYCENSRD 449	
QY 351 -----YEN-DIAMLN-----QAHILPGLNUNAQLF----PPTSMPPPT 387	RESULT 14	
Db 190 KGRSQQPYDPNFPEFTYDGGFMFDRRSPRH--GFEMHARGGFDRMPGPGRPMQS 246	S56329	
QY 388 -----GPPSAMTPYPOEQSETETVHQFIPALSGATIGKQGOHIKOLSPFAGAS- 438	N: Alternative names: nucleic acid-binding protein; protein PCBP-1	
Db 247 RRDYDDMSPRGPRLP-----GRRGR-----GGSRA 274	C:Species: Homo sapiens (man)	
QY 439 L--IKIAPSEPAQAKVAVITGPPDAQFAQGDIYIGKIKEPFVSKEEVKLEAHIRVP 496	C:Date: 15-Feb-1996 # sequence_revision 01-Mar-1996 # text_change 01-Dec-2000	
Db 275 RNLPFLPPPPRPGDR--RGRPHYDMGCGYGRGSFGEDIGGP---VITQQTLPK 326	C:Accession: S56523; S56523; S56523; S41378; S41378;	
QY 497 FAACRVISKGKTVNLQNLISAEEVWVPRDQTDDENQVVKITCHFYAQVAORKIO- 555	R:Killedian, M.; Wang, X.; Liehaber, S.A.	
Db 327 DLAISIISKGQPKIQKIRHEGASKI--DEPLEGSDRITTG-----TODOLONA 377.	EMBO J. 14, 4357-4364, 1995	
QY 556 --LITQVKHQQQKA 568	A:Title: Identification of two KH domain proteins in the alpha-globin mRNA stability	
Db 378 RFLQNSVTKQFSEDYA 393	A:Reference number: S58523; MUID:96016208; PMID:7556077	
RESULT 13	A:Accession: S56529	
T25832 hypothetical protein M01A10.1 - Caenorhabditis elegans	A:Status: preliminary	
C:Species: Caenorhabditis elegans	A:Molecule type: protein	
C:Accession: T25832	R:Leffers, H.; Dejgaard, K.; Celis, J.E.	
R:Scheet, P.	Eur. J. Biochem. 230, 447-453, 1995	
submitted to the EMBL Data Library, February 1997	A:Cross-references: EMBL:X81737; MUD:9460770; PIDN:CAA55016.1; PMID:9460771	
A:Description: The sequence of C. elegans cosmid M01A10.	A:Experimental source: AHA cells (transformed human amnion cells)	
A:Accession: Z20094	A:Note: submitted to the EMBL Data Library, March 1994	
A:Status: preliminary	R:Dasheim, H.C.; Loukianova, T.; Degerdal, A.; Smeilund, E.B.	
A:Molecule type: DNA	Nucleic Acids Res. 22, 959-964, 1994	
A:Residues: 1-680 <SCH>	A:Title: Tissue specific expression and cDNA structure of a human transcript encoding	
A:Cross-references: EMBL:U88174; GSPDB:GN00019; CESP:M01A10.1	A:Reference number: S43489; MUID:94203810; PMID:8152327	
A:Genes: GDB:M01A10.1	A:Accession: S43489	
A:Map position: 1	A:Status: translation not shown	
A:Introns: 1-63/2: 97/3; 121/1: 160/1; 269/3; 411/3; 512/1; 649/47	A:Molecule type: mRNA	
A:Cross-references: EMBL:Z29505	A:Cross-references: EMBL:Z29505	
A:Genetics:	A:Gene: GDB:HNRNPX	
A:Cross-references: EMBL:Z29505	A:Cross-references: GDB:344947	
C:Keywords:	C:Keywords: RNA binding	

Best Local Similarity 22.8%; Pred. No. 4.5e-06; Matches 19; Conservative 70; Mismatches 12; Indels 75; Gaps 12;

Qy 196 LPLRLPPQFGVAGIGKREGATIRNTIKOTSKIDHRKENAGAAEKSITLSPSPTSA 235  
 C:Species: Homo sapiens (man)  
 C:Accession: A44125  
 C:Cross references: GB:M64098; GB:M83789; PID:AAA35962.1; PID:9183892  
 C:Note: sequence extracted from NCBI backbone (NCBIN:106862, NCBIP:106863);  
 C:Superfamily: viginin

Query Match Score 7.0% Best Local Similarity 23.9%; Length 1268;

Matches 95; Conservative 69; Mismatches 153; Indels 81; Gaps 18;

Qy 202 VPTQFVGAGIGKREGATIRNTIKOTSKIDHRKENAGAAEKSITLSPGTSAAKSIL 261  
 C:Species: Homo sapiens (man)  
 C:Accession: A44125  
 C:Cross references: GB:M64098; PID:AAA35962.1; PID:9183892  
 C:Note: sequence extracted from NCBI backbone (NCBIN:106862, NCBIP:106863);  
 C:Superfamily: viginin

Db 157 IPKHEHFRVIGKNEKLQDIEKTAKTKIQPRDD--PSNQKIKTGTGEGIEKARHVI 213  
 C:Species: Homo sapiens (man)  
 C:Accession: A44125  
 C:Cross references: GB:M64098; PID:AAA35962.1; PID:9183892  
 C:Note: sequence extracted from NCBI backbone (NCBIN:106862, NCBIP:106863);  
 C:Superfamily: viginin

Qy 262 EIMKEAQDIFKTFPEIPLTLAHNFVG--RLIGKEGRNKKIEODTKTISPOEL 318  
 C:Species: Homo sapiens (man)  
 C:Accession: A44125  
 C:Cross references: GB:M64098; PID:AAA35962.1; PID:9183892  
 C:Note: sequence extracted from NCBI backbone (NCBIN:106862, NCBIP:106863);  
 C:Superfamily: viginin

Db 214 LI--SAEDRAYERLEVEKAHPFTAPYRLG----TIMETETRINIP--- 260  
 C:Species: Homo sapiens (man)  
 C:Accession: A44125  
 C:Cross references: GB:M64098; PID:AAA35962.1; PID:9183892  
 C:Note: sequence extracted from NCBI backbone (NCBIN:106862, NCBIP:106863);  
 C:Superfamily: viginin

Qy 319 TLYNPFERT-ITVKGVNEYTCAKAEEMKIRESYENDIASMNL---QAHLIPGLNNA 372  
 C:Species: Homo sapiens (man)  
 C:Accession: A44125  
 C:Cross references: GB:M64098; PID:AAA35962.1; PID:9183892  
 C:Note: sequence extracted from NCBI backbone (NCBIN:106862, NCBIP:106863);  
 C:Superfamily: viginin

Db 261 --PSVNRTEEVGEQEQAQAVARI-KKIVEEKKTTIAVEVKSKOHKVISPKN 317  
 C:Species: Homo sapiens (man)  
 C:Accession: A44125  
 C:Cross references: GB:M64098; PID:AAA35962.1; PID:9183892  
 C:Note: sequence extracted from NCBI backbone (NCBIN:106862, NCBIP:106863);  
 C:Superfamily: viginin

Qy 373 LGLSPPTSG---MPPPTS----GPP---SAMTPYPOFESESETVY----HO 410  
 C:Species: Homo sapiens (man)  
 C:Accession: A44125  
 C:Cross references: GB:M64098; PID:AAA35962.1; PID:9183892  
 C:Note: sequence extracted from NCBI backbone (NCBIN:106862, NCBIP:106863);  
 C:Superfamily: viginin

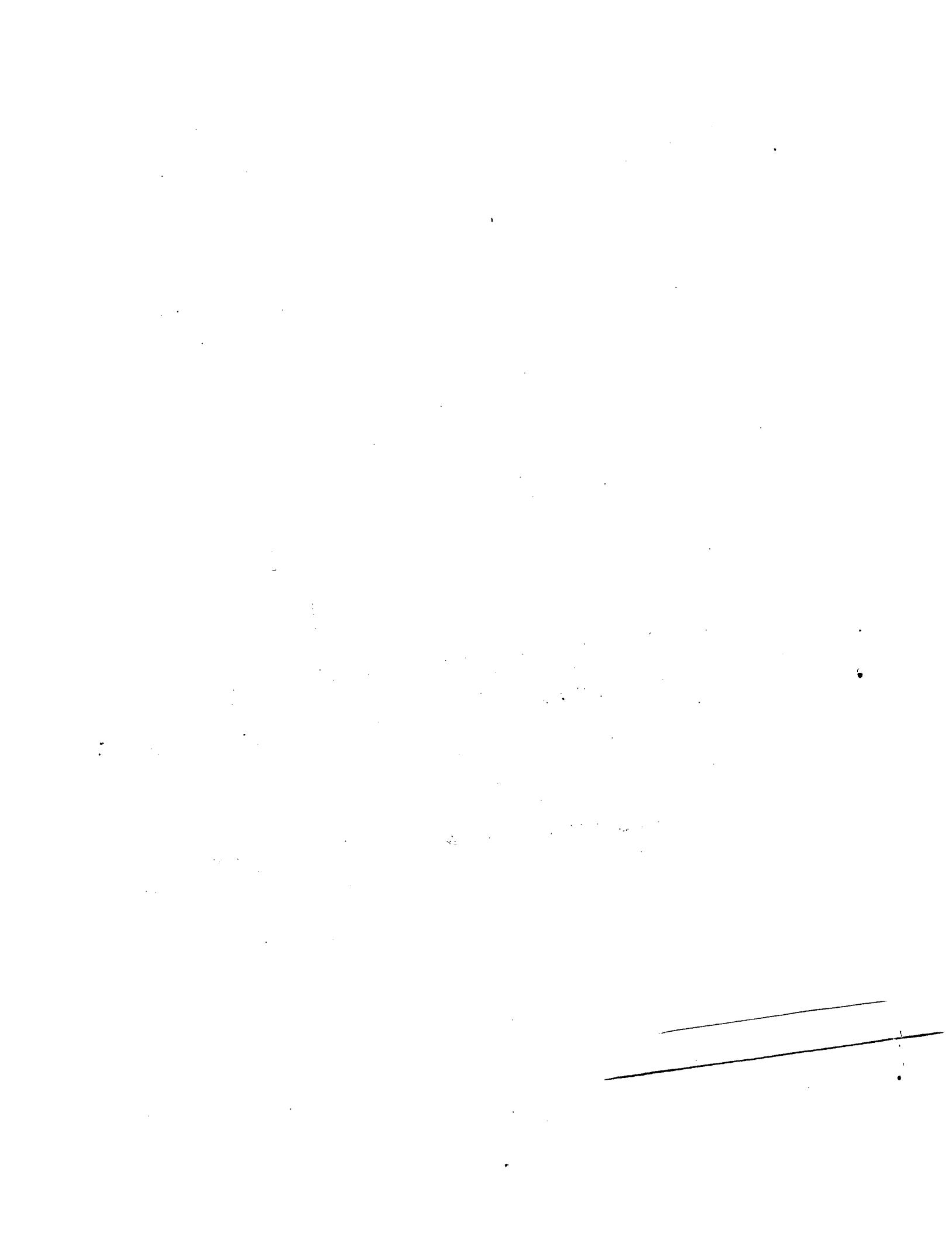
Db 318 LOETLERNPGVSVELPSPSDSISTSTVILREKEPKLQGQATEYVANSPVTSVAAPSMHR 377  
 C:Species: Homo sapiens (man)  
 C:Accession: A44125  
 C:Cross references: GB:M64098; PID:AAA35962.1; PID:9183892  
 C:Note: sequence extracted from NCBI backbone (NCBIN:106862, NCBIP:106863);  
 C:Superfamily: viginin

Qy 411 FIPALSVGATIGKQGHOKIOLSRFAGASIKIAPEADAKVRNVIITGPPEAQFAOGRI 470  
 C:Species: Homo sapiens (man)  
 C:Accession: A44125  
 C:Cross references: GB:M64098; PID:AAA35962.1; PID:9183892  
 C:Note: sequence extracted from NCBI backbone (NCBIN:106862, NCBIP:106863);  
 C:Superfamily: viginin

Db 378 F-----TIGKGONLAKTQ-OMPVHTEFEGED---KITELEGTEVNPQOI 424  
 C:Species: Homo sapiens (man)  
 C:Accession: A44125  
 C:Cross references: GB:M64098; PID:AAA35962.1; PID:9183892  
 C:Note: sequence extracted from NCBI backbone (NCBIN:106862, NCBIP:106863);  
 C:Superfamily: viginin

Qy 471 YGKIKE---ENFVSPKEPVKLEAIIRVPSAAGRIVGGKTVNLQLNSAEVYVPRD 526  
 C:Species: Homo sapiens (man)  
 C:Accession: A44125  
 C:Cross references: GB:M64098; PID:AAA35962.1; PID:9183892  
 C:Note: sequence extracted from NCBI backbone (NCBIN:106862, NCBIP:106863);  
 C:Superfamily: viginin

Db 425 EGMYKDILNMADYVEINIDKHFRHL-----IGKSGANINIKDQYKVSRIV-- 472  
 C:Species: Homo sapiens (man)  
 C:Accession: A44125  
 C:Cross references: GB:M64098; PID:AAA35962.1; PID:9183892  
 C:Note: sequence extracted from NCBI backbone (NCBIN:106862, NCBIP:106863);  
 C:Superfamily: viginin





Qy	1	FVDCPDESVAL 11 1111111111	PRT;	100 AA.	OX NCBI_TAXID=9031; RN [1]; RP SEQUENCE FROM N.A. RX MEDLINE=97220007; PubMed=9121465; RA Ross A. F., Oleynikov Y.S., Kisieluskis E.H., Taneja K.L., Singer R.H.; RT "Characterization of a beta-actin mRNA zipcode-binding protein."; RL Mol. Cell. Biol. 17:2158-2165(1997). DR InterPro; IPR004087; KH_dom. DR InterPro; IPR004088; KH_type_1. DR InterPro; IPR00504; RNA_rec_mot.
Db	41	FVDCPDESVAL 51			
RESULT 2					
Q9D054		PRELIMINARY;			
ID					
AC					
Q9D054;					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 21, Last sequence update)			
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)			
DE	2610101N1Rik	protein.			
IGF-2BP3 OR 2610101N1Rik.					
OS	Mus musculus (Mouse)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Bivalvia; Bivalvia; Rodentia; Sciurognath; Muridae; Murinae; Mus;				
NCBI_TAXID=10059;					
OX					
RN					
SEQUENCE FROM N.A.					
RC	STRAIN=C57BL/6J; TISSUE=EMBRYO;				
RX	MEDLINE=2108560; PubMed=11217851;				
RA	Kawai J., Shingangawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,				
RA	Akara K., Hara A., Fukunishi Y., Konio H., Adachi J., Fukuda S.,				
RA	Aizawa K., Izaya M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,				
RA	Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,				
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,				
RA	Fleischmann W., Gaasterland T., Matsuo Y., Nikaido T., Pesole G., Quackenbush J.,				
RA	Kuehl P., Lewis S., Matsuo C., King B., Kochiwa H., Schriml L.M., Strubli F., Suzuki R., Tomita M., Wagner L., Washio T., Blake J., Boffelli D., Bolunga N., Carninci P., de Bonaldo M.F.,				
RA	Blake J., Boffelli D., Bolunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Garibaldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Rinewald M., Rodriguez T., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shiba Y., Storch K.-F., Suzuki H., Toyoo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y., Functional annotation of a full-length mouse cDNA collection.";				
RT	RL Nature 409:685-690(2001); MGD; MGI:1890359; Igf2bp3.				
DR	InterPro; IPR00504; RNA_rec_mot.				
DR	PF00076; rrm; 1.				
DR	SMART; SM00360; RRM; 1.				
DR	PROSITE; PS0102; RRM; 1.				
SQ	SEQUENCE: 100 AA; 11249 MW; 4D871E37EB9D5466 CRC64;				
Query Match	63.6%; Score 7; DB 11; Length 100;				
Best Local Similarity	100.0%; Pred. No. 0.11;				
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	1	FVDCPDE 7 1111111111	PRT;	576 AA.	Query Match 63.6%; Score 7; DB 4; Length 577;
Db	41	FVDCPDE 4			Best Local Similarity 100.0%; Pred. No. 0.43; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RESULT 3					
O42254		PRELIMINARY;			
ID					
AC					
O42254;					
DT	01-JAN-1998	(TREMBLrel. 05, Created)			
DT	01-JAN-1998	(TREMBLrel. 05, Last sequence update)			
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)			
DE	Zipcode-binding protein.				
CN	ZBP1.				
OS	Gallus gallus (Chicken).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Galliformes; Phasianidae; Phasianae;				
OC	Archosauria; Aves; Neognathae; Galliformes;				
OC	Galins.				

RESULT 5





DR	SMART; SM00322; KH; 4;	RA	Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V., Nester E.W.;
DR	PROSITE; PS50084; KH_TYPE_1; 4.	RA	"Native genome of the natural genetic engineer Agrobacterium tumefaciens C58";
DR	PROSITE; PS50102; RRM; 2.	RT	Science 294:2317-2323(2001).
SQ	SEQUENCE 594 AA; MW; 546CA7BFF0855DD6 CRC64;	RL	[2]
Query Match	Best Local Similarity 63.6%; Score 7; DB 13; Length 594; Matches 7; Conservative 0; Mismatches 0; Indels 0; Caps 0;	RN	SEQUENCE FROM N.A.
Qy	1 FVDCPDE 7	RP	RP MEDLINE=21608551; PubMed=11743194;
Db	1 FVDCPDE 7	RX	RA Goodner B., Hinkley G., Gattung S., Miller N., Blanchard M., Mullin L., Ourolo B., Goldman B.S., Cao Y., Astenazi M., Halling C., Tartachouk O., Epp A., Liu F., Houniel K., Gordon J., Vaudin M., Doughty D., Scott C., Lappas C., Marketz B., Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G., Cielo C., Slater S.
RESULT 10		RT	"Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58."
Q8S3U3	PRELIMINARY; PRT; 225 AA.	RL	Science 294:2323-2328(2001).
ID	Q8S3U3-2002 (TREMBLrel. 21, Created)	DR	AE009072; AA142095.1;
AC		EMBL	AE008038; AAK86891.1;
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)	KW	Complete proteome.
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)	SQ	SEQUENCE 234 AA; 216931DA0BA13758 CRC64;
DE	Prf-like protein (Fragment).	Query Match	Score 6; DB 16; Length 234;
OS	Bukaryota; Viridiplantae; Streptophytina; Embryophytina; Tracheophytina; Magnoliophytina; eudivotsedons; core eudivots; OC	Best Local Similarity 100.0%; Pred. No. 4.1;	
OC	Asteridae; eudicotyledons; core eudicots; OX	Mismatches 0; Indels 0; Gaps 0;	
NCBI_TaxID=4100;	OX	Qy	6 DESWAL 11
RN		DB	113 DESWAL 118
RP	SEQUENCE FROM N.A.	RESULT 12	0981F4 PRELIMINARY; PRT;
RX	MEDLINE=21861904; PubMed=11874570;	ID	QB1F4
RA	"An ED51 orthologue is required for N-mediated resistance against tobacco mosaic virus".	AC	Q94JF4
RT	Pearl J.R., Cook G., Feys B.J., Parker J.E., Baulcombe D.C.; Plant J. 29:569-579(2002)	DT	01-OCT-2001 (TREMBLrel. 18, Created)
RL	EMBL; AF479624; AAL8546.1; -	DT	01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DR		DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)
FT	NON_TER 1	DE	Hypothetical protein m112427.
FT	NON_TER 225	GN	MSL2427.
SQ	SEQUENCE 225 AA; 25674 MW; 1E173D5E15F9860B CRC64;	OS	Rhizobium loti (Mesorhizobium loti).
Query Match	Best Local Similarity 54.5%; Score 6; DB 10; Length 225; Matches 6; Conservative 0; Mismatches 0; Indels 0; Caps 0;	OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
Qy	6 DESWAL 11	OC	Phyllobacteriaceae; Mesorhizobium.
Db	134 DESWAL 139	NCBI_TaxID=381;	OX
RESULT 11		RN	[1]
Q8UGF6	PRELIMINARY; PRT; 234 AA.	RP	SEQUENCE FROM N.A.
ID	Q8UGF6-2002 (TREMBLrel. 21, Created)	RC	SPTRAIN-MAFF10309;
AC		RX	MEDLINE=21082930; PubMed=11214968;
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)	RA	Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S., Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T., Kishida Y., Kiyokawa C., Kohara M., Matsuno A., Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.; RT
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)	RA	"Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizobium loti".
DE	Branches-chain amino acid permease.	RA	DNA Res. 7:331-338(2000).
GN	ATU082 OR ACR_C-2001	DR	AP002999; BAB49562.1; -
OS	Agrobacterium tumefaciens (strain C58 / ATCC 3370).	DR	InterPro; IPR00471; AzIC; 1.
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;	DR	Protein; PF03591; AzIC; 1.
NCBI_TaxID=116299;	OC	KW	Hypothetical Protein; Complete proteome.
RN		SQ	SEQUENCE 244 AA; 25766 MW; EUDCF859FF3CA77F CRC64;
RP	SEQUENCE FROM N.A.	Query Match	Score 6; DB 16; Length 244;
RX	MEDLINE=2160550; PubMed=11743193;	Best Local Similarity 100.0%; Pred. No. 4.2;	
RA	Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P., Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L., Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. St., Chapman P., Clendenning J., Deathridge G., Gillet W., Grant C., Kutyavin T., Levy R., Li M.-J., McClelland E., Palmer J., Palmer A., Raymond C., Rouse G., Saemphumachak C., Wu Z., Romero P., Gordon D., Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M., Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,	Mismatches 0; Indels 0; Gaps 0;	
RA	RESULTS 13	Qy	6 DESWAL 11
RA	QBXOB0	Db	121 DESWAL 126

Best Local Similarity 100.0%; Pred. No. 5.2%; Length 313; Job time : 3.33662 secs

ID	Q8XOB0; Q8XOB0;	PRELIMINARY;	PRT;	250 AA.	Matches 6;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
AC	01-MAR-2002 (TREMBLrel. 20, Created)				QY 1 FVDCPD 6				
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)				Db 11111 299 FVDCPD 304				
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)								
DE	Hypothetical transmembrane protein RSp1376.								
GN	RSP1376 OR RSp02082								
OS	Ralstonia solanacearum (Pseudomonas solanacearum).								
OG	Plasmid megaplasmid.								
OC	Bacteriota; Proteobacteria; beta subdivision; Ralstonia group;								
OC	Ralstonia.								
NCBI_TaxID	305;								
RN	SEQUENCE FROM N.A.								
RC	STRAIN=GM1000;								
RX	MEDLINE=21681879; PubMed=11823852;								
RA	Saïanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,								
RA	Arlat M., Billaut A., Brottet P., Cañus J.C., Cattolico L.,								
RA	Chandler M., Choisne N., Claudel-Renard C., Cunzac S., Demange N.,								
RA	Gaspin C., Lavie M., Moisan A., Robert C., Saunier W., Schiex T.,								
RA	Siguer P., Thébault P., Whalen M., Wincker P., Levy M.,								
RA	Weissenbach J., Boucher C.A.,								
RA	"Genome sequence of the plant pathogen Ralstonia solanacearum.";								
RL	EMBL; AL646084; CAD1B27.1; -.								
DR	InterPro:IPR004471; AzIC; 1.								
KW	Plasmid; Hypothetical protein; Complete proteome.								
SEQUENCE	250 AA; 26490 MW; 685f6FB9089EC2 CRC64;								
SQ									
Query Match	54.58;	Score 6;	DB 16;	Length 250;					
Best Local Similarity	100.08;	Pred. No. 4.3;							
Matches	6;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;				
QY	6 DESNAL 11								
DR	11111 127 DESNAL 132								
RESULT	14								
ID	P96562	PRELIMINARY;	PRT;	310 AA.					
AC	P96562; P96562;								
DT	01-MAY-1997 (TREMBLrel. 03, Created)								
DT	01-MAY-1997 (TREMBLrel. 03, Last sequence update)								
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)								
DE	Putative cytochrome P450 165C2 (Hypothetical hydroxylase D) (Fragment).								
CN	CYP165C2.								
OS	Amycolatopsis orientalis.								
OC	Bacteria; Firmicutes; Actinobacteria; Actinomycetales; Amycolatopsis.								
OC	Pseudonocardiae; Pseudonocardiales; Amycolatopsis.								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=C529-4;								
RA	Soleberg P.J., Matsushima P., Stack D.R., Wilkie S.C., Thompson R.C.,								
RA	Bailey R.H.;								
RL	Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.								
CC	-1 SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.								
DR	EMBL: U84350; AAB49296.1; -.								
DR	HSPP; 000441; LOXA.								
DR	InterPro:IPR01128; Cytochrome_P450.								
DR	pFam:PF00067; PF00067; PF450; 1.								
DR	PROSITE:PS00086; CYTOCHROME_P450; UNKNOWN_1.								
KW	Electron transport; Heme; Hypothetical protein; Membrane;								
KW	Monooxygenase; Oxidoreductase.								
FT	NON_TER 1								
FT	BINDING 260								
SQ	SEQUENCE 310 AA; 34330 MW; 6233C497ED4F8303 CRC64;								
Query Match	54.58;	Score 6;	DB 17;	Length 313;					
Best Local Similarity	100.08;	Pred. No. 5.2;							
Matches	6;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;				
QY	6 DESNAL 11								
DR	11111 127 DESNAL 132								
SEARCH	Completed April 16, 2003, 16:36:41								
Job time :	3.33662 secs								

Best Local Similarity 100.0%; Pred. No. 5.1%; Length 310;

ID	Q9HP19; Q9HP19;	PRELIMINARY;	PRT;	313 AA.	Matches 6;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
AC	01-MAR-2001 (TREMBLrel. 16, Created)				QY 11111 16				
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)				Db 11111 20				
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)				(TREMBLrel. 20, Last annotation update)				
DE	Vog1843C.								
GN	Halobacterium sp. (strain NRC-1).								
OS	Archaea; Euryarchaeota; Halobacteria; Halobacterium.								
OC	Halobacteriaceae; Halobacterium.								
NCBI_TaxID	640911;								
RN	SEQUENCE FROM N.A.								
RP	MEDLINE=2004483; PubMed=11016950;								
RA	Ng W.V., Kennedy S.P., Mahajan G.G., Bergquist B., Pan M.,								
RA	Shukla H.D., Lasky S.P., Baliga N.S., Thorsson V., Sbragno J.,								
RA	Swartzell S., Weir D., Dahl T.A., Welti R., Goo Y.A.,								
RA	Leithauser B., Keller K., Cruz R., Danzon M.J., Hough D.W., Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H., Isenburger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,								
RA	Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D., Elhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S., RT Genome sequence of Halobacterium species NRC-1.								
RL	Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000); DR EMBL: AE003085; AAG2051.1;								
DR	InterPro: IPR000182; GCN5acyltransferf.								
DR	PF00583; Acetyltransferf; 1.								
KW	Complete proteome.								
SQ	SEQUENCE 313 AA; 34330 MW; 6233C497ED4F8303 CRC64;								
Query Match	54.58;	Score 6;	DB 17;	Length 313;					
Best Local Similarity	100.08;	Pred. No. 5.2;							
Matches	6;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;				
QY	6 DESNAL 11								
DR	11111 127 DESNAL 132								
SEARCH	Completed April 16, 2003, 16:36:41								
Job time :	3.33662 secs								

Best Local Similarity 100.0%; Pred. No. 5.1%; Length 310;

ID	P96562	PRELIMINARY;	PRT;	310 AA.	Matches 6;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
AC	01-MAY-1997 (TREMBLrel. 03, Created)				QY 11111 11				
DT	01-MAY-1997 (TREMBLrel. 03, Last sequence update)				Db 186 DESNAL 191				
DE	DESNAL 191								
GN	DESNAL 191								
OS	Amycolatopsis orientalis.								
OC	Actinomycetales; Actinobacteria; Actinomycetidae; Amycolatopsis.								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=C529-4;								
RA	Soleberg P.J., Matsushima P., Stack D.R., Wilkie S.C., Thompson R.C.,								
RA	Bailey R.H.;								
RL	Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.								
CC	-1 SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.								
DR	EMBL: U84350; AAB49296.1; -.								
DR	HSPP; 000441; LOXA.								
DR	InterPro:IPR01128; Cytochrome_P450.								
DR	pFam:PF00067; PF00067; PF450; 1.								
DR	PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.								
KW	Electron transport; Heme; Hypothetical protein; Membrane;								
KW	Monooxygenase; Oxidoreductase.								
FT	NON_TER 1								
FT	BINDING 260								
SQ	SEQUENCE 310 AA; 34330 MW; 0155572AFA694487 CRC64;								
Query Match	54.58;	Score 6;	DB 2;	Length 310;					
Best Local Similarity	100.08;	Pred. No. 5.1;							

GenCore version 5.1.4-p5\_4578  
 Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 24, 2003, 16:11:07 : Search time 38 Seconds  
 (without alignments)  
 3139:507 Million cell updates/sec

Title: US-09-897-778-176  
 Perfect score: 256  
 Sequence: 1 MNKLYIGNLSENAAPSDELS.....YKQHQQQKLQSQQPQSRRK 579

Scoring table: BL0SUM62 Gapext 0.5

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs., 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp Rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteri:\*
- 17: sp\_archeap:\*

012828 homo sapien  
 Q91wjj8 mus musculi  
 Q9bia0 caenorhabdii  
 Q9jxf5 arabidopsis  
 Q9e553 arabidopsis  
 Q96ae4 homo sapien  
 Q9246 homo sapien  
 Q9unw9 homo sapien  
 O43267 homo sapien  
 Q91z82 arabidopsis  
 Q95z29 drosophila  
 Q91kn6 mus musculus  
 Q9bial1 caenorhabdii  
 Q9419 schizosaccharomyces pombe  
 P91612 drosophila  
 Q96sp6 homo sapien  
 Q9ub82 drosophila  
 Q9vh6 drosophila  
 Q9sr13 arabidopsis  
 Q17935 caenorhabdii  
 Q17936 caenorhabdii  
 Q8uvrd9 gallus gallus  
 Q9xJ71 arabidopsis  
 Q9psi1 xenopus laevis  
 Q9asx3 arabidopsis  
 Q9rnk3 arabidopsis  
 Q9ab01 mus musculus  
 Q9mf0 arabidopsis  
 Q8w5c2 oryza sativa

## ALIGNMENTS

RESULT 1  
 ID 000425 PRELIMINARY; PRT; 579 AA.  
 AC 000425; DT 01-JUL-1997 (TREMBLE1. 04, Created)  
 DT 01-JUL-1997 (TREMBLE1. 04, Last sequence update)  
 DT 01-JUN-2002 (TREMBLE1. 21, Last annotation update)  
 DE Putative RNA binding protein KOC (KOC).  
 GN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1] NCBI\_TAXID:9606;  
 RN [2] SEQUENCE FROM N.A.  
 RA Mueller-Pillasch F., Lacher U., Wallrapp C., Et Al.;  
 RL Oncogene 0:0-0(0).  
 RN [2] SEQUENCE FROM N.A.  
 RC TISSUE-PANCREAS;  
 RA Mueller-Pillasch F., Lacher U., Wallrapp C.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL: U97188; AAC35208.1; -;  
 DR InterPro: IPR0004087; KH\_dom.  
 DR InterPro: IPR0004088; KH\_type\_1.  
 DR PFam: PF00013; KH-domain\_4.  
 DR Pfam: P00076; rrm; 2.  
 DR SMART: SM00322; KH; 4;  
 DR PROSITE: P550084; KH\_TYPE\_1; 4.  
 DR PROSITE: P55010; RRM; 2.  
 DR PROSITE: PS00303; RRM\_RNP\_1; FALSE\_NEG.  
 SQ SEQUENCE 579 AA; 63720 MW; AE5C3A8EE3C135CS CRC64;

Query Match 100.0% Score 2956; DB 4; Length 579;  
 Best Local Similarity 100.0% Pred. No. 1.e-187; Mismatches 0; Indels 0; Gaps 0;

## SUMMARIES

Result No	Score	Query	Match	Length	DB	ID	Description
1	2956	100.0	579	4	000425	000425	000425 homo sapien
2	2858	96.7	579	11	Q9cpn8	Q9cpn8 mus musculi	Q9cpn8 mus musculi
3	2483	84.0	593	13	05726	05726 xenopus laevis	05726 xenopus laevis
4	2461.5	83.3	594	13	073932	073932 xenopus laevis	073932 xenopus laevis
5	2261.5	76.5	582	13	09rw80	09rw80 brachydanio	09rw80 brachydanio
6	2219.5	75.1	576	13	042254	042254 gallus gallus	042254 gallus gallus
7	2192	74.2	577	11	088477	088477 mus musculus	088477 mus musculus
8	2181	73.8	577	4	Q9Nq18	Q9Nq18 homo sapien	Q9Nq18 homo sapien
9	1868	63.2	556	4	Q9Y6M1	Q9Y6M1 homo sapien	Q9Y6M1 homo sapien
10	835.5	28.3	558	5	Q9WZ69	Q9WZ69 drosophila	Q9WZ69 drosophila
11	835	28.2	566	5	Q9Nq59	Q9Nq59 drosophila	Q9Nq59 drosophila
12	473	16.0	100	11	Q9D104	Q9D104 mus musculus	Q9D104 mus musculus
13	418.5	14.2	854	5	Q21605	Q21605 caenorhabdii	Q21605 caenorhabdii
14	262.5	8.9	774	10	Q9I128	Q9I128 oryza sativa	Q9I128 oryza sativa
15	260.5	8.8	542	10	Q9S7G1	Q9S7G1	Q9S7G1
16	242.5	8.2	557	5	Q23487	Q23487 caenorhabdii	Q23487 caenorhabdii



Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Rodriguez I., Sakamoto N., Sato K., Schoenbach C., Seya T., Shibusawa Y., Storch K.-F., Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L., Wynnshaw-Boriss A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashiaki Y.	"Functional annotation of a full-length mouse cDNA collection."; Nature 409:685-690(2001).
[2]	SEQUENCE FROM N.A. Mori H., Sakakibara S., Imai T., Nakamura Y., Iijima T., Suzuki A., Yusa H., Takeda M., Okano H.; "Expression of mouse Igf2 mRNA-binding protein 3 and its implications for the developing central nervous system."; J. Neurosci. Res. 0:0 (2001);
RA	EMBL: AK011689; BAB27779_1; EMBL: AB046172; BAB19755_1; MGD: MGFI:1890359; 1972bp3;
RA	DR InterPro: IPRO04087; KH_dom. DR InterPro: IPRO04088; KH_type_1. DR InterPro: IPRO00504; RNA_rec_mot. DR Pfam: PF00013; KH_domain; 4. DR Pfam: PF00076; rtm; 2. SMART: SM00322; KH; 4. DR PROSITE: PS55084; KH_TYPE_1; 4. DR PROSITE: PS51012; RRM; 2. SQ SEQUENCE 579 AA; 63574 MW; CABD9A4355B392B7 CRC64;
RA	Query Match 96.7%; Score 2858; DB 11; Length 579; Best Local Similarity 96.4%; Pred. No. 5.2e-181; Indels 0; Gaps Matches 538; Conservative 10; Mismatches 0;
RA	Qy 1 MNKLYIGNSENAAAPSDELSIFKDAK1PVSGLPFVLYKTYAFVDCPDDESWALKAEALSGK 60 Db 1 MNKLYIGNSDRHDGAPDLSEYFKDAK1PVAQPFVLYKTYAFVDCPDDESWALKAEALSGK 60
RA	Qy 61 IELHGKPTEVERSVPKRQIRKLQTNRNPIPHLQWEVLDLIVQGVVESCEOVNNTDSETA 120 Db 61 MELHGKPTEVERSVPKRQIRKLQTNRNPIPHLQWEVLDLIVQGVVESCEOVNNTDSETA 120
RA	Qy 121 VVNTYSSKQDAQAKLKGFLQLENFTLKVAYIDEMAAQQNQLQQPQRGRGLGQRGSS 180 Db 121 VVNTYSSKQDAQAKLKGFLQLENFTLKVAYIDETAAQQNNSPQLRGRGPQRGGSS 180
RA	Qy 181 RQSPGSYSKQKCDPLRLVPTQFGATIGKEGATIRNITKTQSKDVHKEKGAA 240 Db 181 RQSPGSYSKQKCDPLRLVPTQFGAIIKGKEGATIRNITKTQSKDVHKEKGAA 240
RA	Qy 241 EKSITLSTPEGTSAAACKSLLIMHKEAQD1KFTEEPKTLAANNFGRGLGEGRNLK 300 Db 241 EKSITLSTPEGTSAAACKSLLIMHKEAQD1KFTEEPKTLAANNFGRGLGEGRNLK 300
RA	Qy 301 KIEQDTDKITISPLQELTLYNPERITVKSYNECAKAEFIMKKIRESYENDIASML 360 Db 301 KIEQDTDKITISPLQELTLYNPERITVKSYNECAKAEFIMKKIRESYENDIASNL 360
RA	Qy 361 QAHJLPGLNNAQAGLEPPTSGMPPTSGPPSAMTPPYQPEQESETETYHQFIPALSYGAI 420 Db 361 QAHJLPGLNNAQAGLEPPTSGMPPTSGPPSAMTPPYQPEQESETETVHFLPALSYGAI 420
RA	Qy 421 IGKQGQHIIKOLSRPAGASIKTAAEPADAKYRMVITGPBQAQKAQGRIGYTKKEENF 480 Db 421 IGKQGQHIIKOLSRPAGASIKTAAEPADAKYRMVITGPBQAQKAQGRIGYTKKEENF 480
RA	Qy 481 SPKEEVKLEAHIRYPSFAAGRYIGKGGKTYNELQNLSSAVVVPRDQTEDNDQVVVKIT 540 Db 481 SPKEEVKLEAHIRYPSFAAGRYIGKGGKTYNELQNLSSAVVVPRDQTEDNDQVVVKIT 540
RA	Qy 541 GHFYACQVAQRKIQEILTQVKOHQQOKALQSGPQSRK 579 Db 541 GHFYACQVAQRKIQEILTQVKOHQQOKALQSGPQSRK 579

RESULT	2	PRY
29CPNB	Q9CPNB	PRY
QAC	Q9CPNB;	
DTT	01-JUN-2001 (T)	
DTT	01-JUN-2001 (T)	
DE	10 days embryo	
DE	clone: 2610036B	
DE	3).	
DE	IGF2BP3 OR 2611	
COS	Mus musculus (	
OC	Eukaryota; Met-	
OX	Mammalia; Eutho-	
OX	NCBI_TAXID=100	
RRN	[1]	
	SEQUENCE FROM:	
	STRAN= C57BL/6	
	MEDLINE=210856	
	RX	
	Kawai J., Shinohara T., Arakawa T., Aizawa K., Saito T., Kadono K., Fleischmann W., Kuehnl P., Lewis RA, Schrimm L.M., Sakai K., Okuda RA, Blake J., Boffo RA, Brownstein M.J., Sacchi S.	

RESULT  
057526  
ID  
AC  
7

